GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 08:46:36 ; Search time 2644 Seconds
               (without alignments)
               6098.188 Million cell updates/sec

Title: US-09-972-916B-5
Perfect score: 372
Sequence: 1 tacactggggccagagtcc...........ttccgctactagctagccgc 372

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151699S residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
        1: gb_ba:*
        2: gb_htg:*
        3: gb_in:*
        4: gb_om:*
        5: gb_ov:*
        6: gb_pat:*
        7: gb_ph:*
        8: gb_pl:*
        9: gb_pr:*
       10: gb_ro:*
       11: gb_sts:*
       12: gb_sy:*
       13: gb_un:*
       14: gb_vi:*
       15: gb_ba:*
       16: gb_fun:*
       17: gb_hum:*
       18: em_in:*
       19: em_mu:*
       20: em_om:*
       21: em_or:*
       22: em_ov:*
       23: em_pat:*
       24: em_ph:*
       25: em_pl:*
       26: em_ro:*
       26: em_sts:*
       27: em_un:*
       28: em_vi:*
       29: em_vrt:*
       30: em_hum:*
       31: em_htg_inv:*
       32: em_htg_other:*
       33: em_htg_mus:*
       34: em_htg_pln:*
       35: em_htg_rod:*
       36: em_htg_mam:*
       37: em_htg_vrt:*
       38: em_sy:*
       39: em_htgo_hum:*
       40: em_htgo_mus:*
       41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	214.4	57.6	1181	10	RATILGFZ	M84484 Rattus norv
2	214.4	57.6	185148	2	AC136382	AC136382 Rattus no
3	188.8	50.8	1363	10	MMILGF	X67493 M.musculus
4	188.8	50.8	192843	2	AL607124	AL607124 Mouse DNA
5	98.2	26.4	1569	10	BC013345	BC013345 Mus muscu
6	76.4	20.5	1500	6	AX401932	AX401932 Sequence
7	76.4	20.5	1500	6	AX827271	AX827271 Sequence
8	76.4	20.5	1500	10	RATIGFB	M58634 Rat IGF bin
9	76.4	20.5	5001	6	RATIGFB	AX163782 Sequence
10	76.4	20.5	5001	10	RATIGFBA	AX163782 Sequence
11	64	17.2	3886	9	AY093345	AY093345 Papio anu
12	63.8	17.2	6128	9	AX409747	AX409747 Sequence
13	63.8	17.2	6128	9	HUMIGFBP1A	HUMIGFBP1A Human insul
14	63.8	17.2	6128	11	G19994	G19994 Rat insul
15	63.8	17.2	9082	9	AY434089	AY434089 Homo sapi
16	63.8	17.2	69887	9	AC091524	AC091524 Homo sapi
c 17	62.2	16.7	141539	9	AC146152	AC146152 Homo sapi
18	62.2	16.7	182079	2	AC146152	AC146152 Pan trogl
19	62.2	16.7	448	11	AC146117	AC146117 Pan trogl
20	60.8	16.3	194	6	G67139	G67139 IGFBP1 Xl.1
21	55.2	14.8	194	6	A57715	A57715 Sequence 1
22	55.2	14.8	13011	6	AR175909	AR175909 Sequence 1
23	55.2	14.8	13011	6	E14395	E14395 gDNA encodi
24	55.2	14.8	13011	10	AX827302	AX827302 Sequence
25	55.2	14.8	13011	10	RNLPKG	X05684 Rat L-PK ge
26	54.4	14.6	895	9	HUMIGFBPB1	M23592 Human insul
27	52	14.0	539	11	G67173	G67173 IGFBP1 5'A
c 28	47.4	12.7	2717	10	RATPKRL1	M17008 Rat pyruvat
c 29	47.4	12.7	231241	2	AC097039	AC097039 Rattus no
c 30	47.2	12.7	199854	2	AC132327	AC132327 Mus muscu
c 31	44.6	12.0	184889	9	AL442125	AL442125 Human DNA
32	44.6	12.0	256781	2	AC097952	AC097952 Rattus no
33	44.6	12.0	259329	2	AC105670	AC105670 Rattus no
c 34	44.4	11.9	1480	10	RATIGFBP1	M89791 Rattus norv
c 35	44.2	11.9	206515	10	AC140332	AC140332 Mus muscu
c 36	42.6	11.5	43295	9	AC006177	AC006177 Homo sapi
c 37	42.6	11.5	53370	9	AL592071	AL592071 Human DNA
c 38	42.6	11.5	186314	10	AC003694	AC003694 Mus muscu
39	42.4	11.4	185452	2	AC114566	AC114566 Mus muscu
c 40	42.4	11.4	187942	2	AC116502	AC116502 Mus muscu
41	41.4	11.1	200340	9	AC100797	AC100797 Homo sapi
42	41	11.0	81704	9	AC110299	AC110299 Homo sapi
43	41	11.0	151700	9	AC133528	AC133528 Homo sapi
44	41	11.0	182444	2	AC116481	AC116481 Mus muscu
45	40.6	10.9	125020	9	AF429315	AF429315 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS   RATILGFZ   1181 bp   DNA   linear   ROD 27-APR-1993
DEFINITION Rattus norvegicus insulin-like growth factor gene fragment.
ACCESSION   M84484
VERSION   M84484.1 GI:204927
KEYWORDS   insulin-like growth factor.
SOURCE   Rattus norvegicus
ORGANISM   Rattus norvegicus (Norway rat)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 1181)
AUTHORS   Unterman,T.G., Lacson,R.G., McGary,E., Whalen,C. and Goswami,R.G.
JOURNAL   Biochem. Biophys. Res. Commun. (1991) In press

COMMENT Original source text: Rattus norvegicus (strain Sprague-Dawley)
    male adult liver DNA.
FEATURES   Location/Qualifiers
 source   1..1181
      /organism="Rattus norvegicus"
      /mol_type="genomic DNA"
      /strain="Sprague-Dawley"
      /db_xref="taxon:10116"
      /sex="male"
      /tissue_type="liver"
      /dev_stage="adult"
ORIGIN

Query Match
 Best Local Similarity 57.6%;
 Matches 218; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 144 TGCGCCCATGTCACAAGCAAAACAAAACTTATTTATTTGAACACGGGGGATCCTAGCACGCTGCC 203
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 813 TGTGCGGAGCTCACAACAAAACAAAACTTATTTATTTGAACACGGGGGATCCTAGCACGCTGCC 872

Qy 204 CTGACAAATCAATTAACCCGTGCTGCCGGAGCCAGCCCTTCATAAGGCCCTGGGTATGGCCAG 263
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 873 CTGACAAATCAATTAACCCGTGCTGCCGGAGCCAGCCCTTCATAAGGCCCTGGGTATGGCCAG 932

Qy 264 CCAGCACGGTCCACTGCCCGCCGCGAGACAACAAACCCGCGAGCGATTGAACACTGCACACGG 323
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 933 CCAGCACGGTCCACTGCCCGCCGCGAGACAACAAACCCGCGAGCGATTGAACACTGCACACGG 992

Qy 324 CCATCTGCCCAGAGAGCTGTGACCACCACACTTCCGCTACTAGCTA 367
     ||||||||||||||||||||||||||||||||||||||||||||||
Db 993 CCATCTGCCCAGAGAGCTGTGACCACCACACTTCCGCTACTATCTA 1036

RESULT 2
LOCUS   AC136382 185148 bp DNA   linear HTG 01-NOV-2002
DEFINITION Rattus norvegicus clone CH230-97018, *** SEQUENCING IN PROGRESS
    ***, 63 unordered pieces.
ACCESSION AC136382
VERSION AC136382.1 GI:24462257
KEYWORDS HTG; HTGS_PHASE1.
SOURCE   Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
    Rattus.
REFERENCE 1 (bases 1 to 185148)
 AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
    Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
    Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
    Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
    Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
    Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
    Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
    Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
    Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
    Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
    Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
    Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
    Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,C.,
    Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
    Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
    Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
    Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
    Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
    Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
    Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
    Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
    Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,J., Kovar,C.,
    Kowis,C., Kraft,C.I., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
    Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lu,X., Ma,J.,
    Lorensuhewa,L., Louiseged,H., Lozado,R.J., Lu,Y., Lu,X., Ma,J.,

    Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
    Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
    Mawhiney,S., Mcleod,M., Mcneill,T., Meenen,E., Milosavljevic,A.,
    Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,
    Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D.,
    Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G.,
    Olarinpunagoon,A., Pai,S., Parks,K., Pasternak,S., Paul,H.,
    Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A.,
    Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E.,
    Reeves,K., Regier,M., Reilly,R., Reilly,B., Reilly,M., Ren,Y.,
    Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A.,
    Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S.,
    Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A.,
    Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E.,
    Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A.,
    Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,
    Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D.,
    Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J.,
    Warren,R., Wei,X., White,F., Williams,J., Willson,R., Wleczyk,R.,
    Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,
    Yen,J., Yoon,L., Yu,Y., Zhang,J., Zhou,J., Zhou,X.,
    Zhao,S., Dunn,D., von Niederhausern,A., Weise,R., Smith,D.R.,
    Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
 TITLE   Direct Submission
 JOURNAL Unpublished
REFERENCE 2 (bases 1 to 185148)
 AUTHORS Rat Genome Sequencing Consortium.
 TITLE   Direct Submission
 JOURNAL Submitted (01-NOV-2002) Human Genome Sequencing Center, Department
    of Molecular and Human Genetics, Baylor College of Medicine, One
    Baylor Plaza, Houston, TX 77030, USA
COMMENT   Center: Baylor College of Medicine
    Center code: BCM
    Web site: http://www.hgsc.bcm.tmc.edu/
    Contact: hgsc-help@bcm.tmc.edu
    --------------- Project Information ---------------
    Center project name: KDGS
    Center clone name: CH230-97018
    --------------- Summary Statistics ---------------
    Sequencing vector: plasmid;
    Chemistry: Dye-terminator Big Dye: 100% of reads
    Assembly program: Phrap; version 0.990329
    Consensus quality: 135613 bases at least Q40
    Consensus quality: 140849 bases at least Q30
    Consensus quality: 145680 bases at least Q20
    ----
    * NOTE: Estimated insert size may differ from sequence length
    * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
    * NOTE: This is a 'working draft' sequence. It currently
    * consists of 63 contigs.The true order of the pieces
    * is not known and their order in this sequence record is
    * arbitrary. Gaps between the contigs are represented as
    * runs of N, but the exact sizes of the gaps are unknown.
    * This record will be updated with the finished sequence
    * as soon as it is available and the accession number will
    * be preserved.
    *
    * 1 1010: contig of 1010 bp in length
    * 1011 1110: gap of unknown length
    * 1111 2608: contig of 1498 bp in length
    * 2609 2708: gap of unknown length
    * 2709 3915: contig of 1207 bp in length
    * 3916 4015: gap of unknown length
    * 4016 5547: contig of 1532 bp in length
    * 5548 5647: gap of unknown length
    * 5648 7119: contig of 1472 bp in length
    * 7120 7219: gap of unknown length
    * 7220 8885: contig of 1666 bp in length
    * 8886 8985: gap of unknown length
    * 8986 10258: contig of 1273 bp in length
    * 10259 10358: gap of unknown length
    * 10359 11626: contig of 1268 bp in length
    * 11627 11726: gap of unknown length

```
*  11727  13488: contig of 1762 bp in length
*  13489  13588: gap of unknown length
*  13589  15122: contig of 1534 bp in length
*  15123  15222: gap of unknown length
*  15223  16946: contig of 1724 bp in length
*  16947  17046: gap of unknown length
*  17047  18164: contig of 1118 bp in length
*  18165  18264: gap of unknown length
*  18265  19678: contig of 1414 bp in length
*  19679  19778: gap of unknown length
*  19779  21687: contig of 1909 bp in length
*  21688  21787: gap of unknown length
*  21788  22892: contig of 1105 bp in length
*  22893  22992: gap of unknown length
*  22993  24336: contig of 1344 bp in length
*  24337  24436: gap of unknown length
*  24437  26617: contig of 2181 bp in length
*  26618  26717: gap of unknown length
*  26718  28577: contig of 1860 bp in length
*  28578  28677: gap of unknown length
*  28678  30823: contig of 2146 bp in length
*  30824  30923: gap of unknown length
*  30924  32266: contig of 1343 bp in length
*  32267  32366: gap of unknown length
*  32367  33894: contig of 1528 bp in length
*  33895  33994: gap of unknown length
*  33995  35373: contig of 1379 bp in length
*  35374  35473: gap of unknown length
*  35474  37295: contig of 1822 bp in length
*  37296  37395: gap of unknown length
*  37396  38420: contig of 1025 bp in length
*  38421  38520: gap of unknown length
*  38521  39974: contig of 1454 bp in length
*  39975  40074: gap of unknown length
*  40075  42283: contig of 2209 bp in length
*  42284  42383: gap of unknown length
*  42384  43721: contig of 1338 bp in length
*  43722  43821: gap of unknown length
*  43822  46538: contig of 2717 bp in length
*  46539  46638: gap of unknown length
*  46639  48621: contig of 1983 bp in length
*  48622  48721: gap of unknown length
*  48722  50602: contig of 1881 bp in length
*  50603  50702: gap of unknown length
*  50703  53841: contig of 3139 bp in length
*  53842  53941: gap of unknown length
*  53942  56195: contig of 2254 bp in length
*  56196  56295: gap of unknown length
*  56296  58949: contig of 2654 bp in length
*  58950  59049: gap of unknown length
*  59050  61031: contig of 1983 bp in length
*  61032  61131: gap of unknown length
*  61132  62941: contig of 1810 bp in length
*  62942  63042: gap of unknown length
*  63043  66045: contig of 3003 bp in length
*  66046  66145: gap of unknown length
*  66146  68197: contig of 2052 bp in length
*  68198  68297: gap of unknown length
*  68298  72235: contig of 3938 bp in length
*  72236  72335: gap of unknown length
*  72336  75803: contig of 3468 bp in length
*  75804  75903: gap of unknown length
*  75904  78355: contig of 2452 bp in length
*  78356  78455: gap of unknown length
*  78456  81458: contig of 3003 bp in length
*  81459  81558: gap of unknown length
*  81559  84673: contig of 3115 bp in length
*  84674  84773: gap of unknown length
*  84774  87739: contig of 2966 bp in length
*  87840  87839: gap of unknown length
*  87840  91535: contig of 3696 bp in length
*  91536  91635: gap of unknown length
*  91636  95522: contig of 3887 bp in length
*  95523  95622: gap of unknown length
*  95623  98730: contig of 3108 bp in length
*  98731  98830: gap of unknown length
*  98831  101618: contig of 2788 bp in length
*  101619 101718: gap of unknown length
*  101719 105535: contig of 3817 bp in length
*  105536 105635: gap of unknown length
*  105636 109393: contig of 3758 bp in length
*  109394 109493: gap of unknown length
*  109494 113564: contig of 4071 bp in length
*  113565 113664: gap of unknown length
*  113665 117630: contig of 3966 bp in length
*  117631 117730: gap of unknown length
*  117731 122502: contig of 4772 bp in length
*  122503 122602: gap of unknown length
*  122603 126644: contig of 4042 bp in length
*  126645 126744: gap of unknown length
```

```
RESULT 3
MMILGF
LOCUS       MMILGF       1363 bp    DNA     linear   ROD 01-AUG-1996
DEFINITION  M.musculus gene for insulin-like growth factor binding protein-1.
ACCESSION   X67493
VERSION     X67493.1  GI:52699
KEYWORDS    insulin-like growth factor binding protein-1.
SOURCE      Mus musculus
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1363)
AUTHORS     Lee,J., Greenbaum,L., Haber,B.A., Nagle,D., Lee,V., Miles,V.,
            Mohn,K.L., Bucan,M. and Taub,R.
TITLE       Structure and localization of the IGFBP-1 gene and its expression
            during liver regeneration
JOURNAL     Hepatology 19 (3), 656-665 (1994)
MEDLINE     94164648
PUBMED      7509771
REFERENCE   2  (bases 1 to 1363)
AUTHORS     Mohn,K.L., Waddle,J.R. and Taub,R.
TITLE       Comparison of mouse and human IGFBP-1 genes reveals a potential
            insulin-responsive sequence and conservation of all intron/exon
            boundaries
JOURNAL     Nucleic Acids Res.
REFERENCE   3  (bases 1 to 1363)
AUTHORS     Taub,R.A.
TITLE       Direct Submission
JOURNAL     Submitted (23-JUL-1992) R.A. Taub, Univ. of Pennsylvania, Howard
            Hughes Medical Institute, Clinical Research Bldg., Room 475, 422
            Curie Boulevard, Philadelphia, PA 19104-6145, USA
FEATURES             Location/Qualifiers
     source          1..1363
                     /organism="Mus musculus"
```

```
Query Match             57.6%;  Score 214.4;  DB 2;  Length 185148;
Best Local Similarity   97.3%;  Pred. No. 1.6e-51;
Matches  218;  Conservative  0;  Mismatches  6;  Indels  0;  Gaps  0;

Qy    144 TGCGCCCATGTCACAAGCAAAACAAAACTTATTTGAACACAGGGGATCCTAGCCAGGCTGCC 203
          |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  99243 TGTGCGCGGAGCTCACAAGCAAAAGCAAAACAAAACTTATTTGAACACACGGGGATCCTAGCCAGGCTGCC 99302

Qy    204 CTGACAAATCATTAACCCGTGCTGCCCGAGCCAGCCCCTTCATAAGGCCCTGGGTATGGCCAG 263
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  99303 CTGACAAATCATTAACCCGTGCTGCCCGAGCCAGCCCCTTCATAAGGCCCTGGGTATGGCCAG 99362

Qy    264 CCAGCATGGTCCACTGCCCCGCCCGAGACACAAAACCCGAGCGCATTGAACACTGCCACACGG 323
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  99363 CCAGCATGGTCCACTGCCCCGCCCGAGACACAAAACCCGAGCGCATTGAACACTGCCACACGG 99422

Qy    324 CCATCTGCCCAGAGAGCTGTGACCACCACTTCCGCTACTAGCTA 367
          |||||||||||||||||||||||||||||||||||||||||||
Db  99423 CCATCTGCCCAGAGAGCTGTGACCACCACTTCCGCTACTATCTA 99466
```

COMMENT   During sequence assembly this sequence data is compared from overlapping clones.
            On Apr 12, 2002 this sequence version replaced gi:19847866.
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            Submitted (11-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
REFERENCE   1
AUTHORS   Oliver.K.
TITLE   Direct Submission
JOURNAL   Submitted...

```
            /mol_type="genomic DNA"
            /strain="NIH Swiss"
            /db_xref="taxon:10090"
            /clone="JW2.1"
            /cell_line="NIH 3T3"
            /cell_type="fibroblast"
            /clone_lib="genomic, lambda FIX II"
            /dev_stage="embryo"
misc_feature     500..508
            /note="AP-2 consensus site"
misc_feature     692..706
            /note="insulin-responsive element"
CAAT_signal      720..741
            /note="Caat box with APF,HNF and NF-E1 consensus
            sequences"
TATA_signal      764..768
exon             792..1336
            /number=1
misc_feature     792
            /note="mRNA cap site"
CDS              964..>1336
            /codon_start=1
            /product="insulin-like growth factor binding protein-1"
            /protein_id="CAA47832.1"
            /db_xref="GI:52700"
            /db_xref="GOA:P47876"
            /db_xref="SWISS-PROT:P47876"
            /translation="MPEFITVVSWPFLILLSFQIGVAAGAPQVWHCAPCTAERLGLCP
            PVPASCPEISRPAGCGCCPTCALPMGAACGVATARCAQGLSCRALPGERPLHALTRG
            QGASLPEPAAPATSTLFSSQHE"
ORIGIN
```

RESULT 4

LOCUS   AL607124   192843 bp   DNA   linear   ROD 11-APR-2002
DEFINITION Mouse DNA sequence from clone RP23-20C9 on chromosome 11, complete
            sequence.
ACCESSION   AL607124
VERSION   AL607124.15 GI:20145926
KEYWORDS   HTG.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Query Match   50.8%; Score 188.8; DB 10; Length 1363;
Best Local Similarity   90.2%; Pred. No. 3.6e-44;
Matches 202; Conservative   0; Mismatches 22; Indels   0; Gaps   0;

```
Qy    144 TGCGCCCATGTCACAAGCAAAACAAAACTTATTTGAACACGGGGATCCTAGCACGCTGCC 203
          ||||||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db    667 TGTGTAGAGCTCACAAGCAAAACAAAACTTATTTGAACACTGGGGTCCTAGCACGGCTGCC 726

Qy    204 CTGACAAATCATTAACCCGTGCTGCCCAGCCCTTCATAAGGCCCTGGGTATGGCCAG 263
          |||||||||||||||| ||||| || |||||||||||||||| ||||||  |  |||
Db    727 CTGACAAATCATTAACCTGTGCCGCACAGCCCTTCATAAGGCTCTGGGTTACGACCAG 786

Qy    264 CCAGCANGGTCCACTGCCCCGCCGGAGACACAAACCCAGGCCATTGAACACTGCCACACGG 323
          |||||| |||||||||| ||||||||||||| |||||||||||||||||||||||||||||
Db    787 CCAGCATGGTCCACTGCCCGCCGGAGACACACCCAGGCCATTGAACACTGCCACACGG 846

Qy    324 CCATCTGCCCAGAGAGCTGTGACCACCACTTCCGCTACTAGCTA 367
          || ||| ||||||||||||||||||||||| ||| ||||| ||||
Db    847 CCGTCTCGCCCAGAGAGCTGTGACCACCATTGCCACTACTATCTA 890
```

FEATURES   Location/Qualifiers
     source   1..192843
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="11"
                     /clone="RP23-20C9"
                     /clone_lib="RPCI-23"
VECTOR: pBACe3.6.

ORIGIN

Query Match   50.8%; Score 188.8; DB 10; Length 192843;
Best Local Similarity   90.2%; Pred. No. 4.6e-44;
Matches 202; Conservative   0; Mismatches 22; Indels   0; Gaps   0;

```
Qy    144 TGCGCCCATGTCACAAGCAAAACAAAACTTATTTGAACACGGGGATCCTAGCACGCTGCC 203
          ||||||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db   2419 TGTGTAGAGCTCACAAGCAAAACAAAACTTATTTGAACACTGGGGTCCTAGCACGGCTGCC 2478

Qy    204 CTGACAAATCATTAACCCGTGCTGCCCAGCCCTTCATAAGGCCCTGGGTATGGCCAG 263
          |||||||||||||||| ||||| || |||||||||||||||| ||||||  |  |||
Db   2479 CTGACAAATCATTAACCTGTGCCGCACAGCCCTTCATAAGGCTCTGGGTTACGACCAG 2538

Qy    264 CCAGCANGGTCCACTGCCCCGCCGGAGACACAAACCCAGGCCATTGAACACTGCCACACGG 323
          |||||| |||||||||| ||||||||||||| |||||||||||||||||||||||||||||
Db   2539 CCAGCATGGTCCACTGCCCGCCGGAGACACACCCAGGCCATTGAACACTGCCACACGG 2598

Qy    324 CCATCTGCCCAGAGAGCTGTGACCACCACTTCCGCTACTAGCTA 367
          || ||| ||||||||||||||||||||||| ||| ||||| ||||
Db   2599 CCGTCTCGCCCAGAGAGCTGTGACCACCATTGCCACTACTATCTA 2642
```

RESULT 5

LOCUS   BC013345   1569 bp   mRNA   linear   ROD 12-NOV-2003
DEFINITION Mus musculus insulin-like growth factor binding protein 1, mRNA
            (cDNA clone MGC:14075 IMAGE:4161889), complete cds.
ACCESSION   BC013345
VERSION   BC013345.1 GI:15426482
KEYWORDS   MGC.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1569)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

---

Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e. phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; Sw:,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-20C9 is
from the RPCI-23 Mouse PAC Library.
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm

REFERENCE   2 (bases 1 to 1569)
AUTHORS   Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (31-AUG-2001) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Institute for Systems Biology
             http://www.systemsbiology.org
             contact: amadan@systemsbiology.org
             Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
             Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

             Clone distribution: MGC clone distribution information can be found
             through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
             Series: IRAX Plate: 18 Row: i Column: 20
             This clone was selected for full length sequencing because it
             passed the following selection criteria: Hexamer frequency ORF
             analysis, Similarity but not identity to protein.
FEATURES   Location/Qualifiers
     source   1..1569
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="MGC:14075 IMAGE:4161889"
                     /tissue_type="Liver, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Li9"
                     /lab_host="DH10B"
                     /note="vector: pCMV-SPORT6"
     gene   1..1569
                     /gene="Igfbp1"
                     /note="synonym: IGFBP-1"
                     /db_xref="LocusID:16006"
                     /db_xref="MGI:96435"
     CDS   204..1022
                     /codon_start=1
                     /product="insulin-like growth factor binding protein 1"
                     /protein_id="AAH13345.1"
                     /db_xref="GI:15426483"
                     /translation="MPEFLTVSWPFLILSFGIGVAAGAPQPWHCAPCTAERLGLCP
                     PVPASCPEISRPAGCGCCPTCALPMGAACGVATARCAQGLSCRALPGEPRPLHALTRG
                     QGACVPPAAPAIPSTLSSSQHEEAKAVVSADELSESPEMTEEQLLDSFHLMAPSRED
                     QPILWNAISTYSSMRAREIADLKKWKEPCQRELYKVLERLAAAQQKAGDEIYKFYLPN
                     CNKNGVYHSKQCETSLDGEAGLCWCVYPWSGKKIPGSLETRGDPNCHQYFNVHN"
     misc_feature   291..527
                     /note="IB; Region: Insulin growth factor-binding protein
                     homologues"
                     /db_xref="CDD:smart00121"
     misc_feature   768..995
                     /note="thyroglobulin_1; Region: Thyroglobulin type-1

---

repeat. Thyroglobulin type 1 repeats are thought to be
involved in the control of proteolytic degradation. The
domain usually contains six conserved cysteines. These
form three disulphide bridges. Cysteines 1 pairs with 2, 3
with 4 and 5 with 6"
                     /db_xref="CDD:pfam00086"
ORIGIN

RESULT 6
LOCUS   AX401932   1500 bp   DNA   linear   PAT 06-JUN-2002
DEFINITION Sequence 1608 from Patent WO0210453.
ACCESSION   AX401932
VERSION   AX401932.1 GI:21338112
KEYWORDS   .
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Mendrick,D., Porter,M.W., Johnson,K.R., Castle,A.L. and
            Elashoff,M.R.
  TITLE   Molecular toxicology modeling
  JOURNAL   Patent: WO 0210453-A 1608 07-FEB-2002;
            Gene Logic, Inc. (US)
FEATURES   Location/Qualifiers
     source   1..1500
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
                     /note="EMBL/GenBank Accession No. NM_013144"
ORIGIN

Query Match   26.4%; Score 98.2; DB 10; Length 1569;
Best Local Similarity   92.8%; Pred. No. 1e-17;
Matches 103; Conservative   0; Mismatches   8; Indels   0; Gaps   0;

Qy   257 TGGGCCAGCCAGCCATGGTCCCACTCGCCGCCGCCAGACACAAAACCCAGGGAGCAGCATTGAACACTG 316
          ||||||||||||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||||||
Db   20 TGGGCCAGCCAGCCATGGTCCCACTCGCCGCCGCCAGACACAAAACCCAGGGAGCAGCATTGAACACTG 79

Qy   317 CACACGGCCCATCTGCCCAGAGAGCTGTGACCACCACTTCCGCTACTAGCTA 367
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 CACACGGCCCCTCTGCCCAGAGAGAGCTGTGACCACCATTGCCACTACTATCTA 130

RESULT 7
LOCUS   AX827271   1500 bp   DNA   linear   PAT 12-DEC-2003
DEFINITION Sequence 5 from Patent EP1344834.
ACCESSION   AX827271
VERSION   AX827271.1 GI:39837360
KEYWORDS   .
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1

Query Match   20.5%; Score 76.4; DB 6; Length 1500;
Best Local Similarity   98.7%; Pred. No. 2.3e-11;
Matches 77; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   290 CACAAAACCCAGCGAGCGAGCATTGAACAACTGCACACGGCCCATCTGCCCAGAGAGGAGCTGTGACCAC 349
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8 CACAAACCCAGCGAGCGAGCATTGAACACTGCACACGGCCCATCTGCCCAGAGAGGAGCTGTGACCAC 67

Qy   350 CACTTCCGCTACTAGCTA 367
          ||||||||||||||||||
Db   68 CACTTCCGCTACTATCTA 85

REFERENCE   1

AUTHORS   Boess,F., Suter-Dick,L. and Wolf,D.
TITLE   Methods for the toxicity prediction of a compound
JOURNAL   Patent: EP 1344834-A 5 17-SEP-2003;
          P. HOFFMANN-LA ROCHE AG (CH)
FEATURES Location/Qualifiers
  source 1..1500
          /organism="Rattus norvegicus"
          /mol_type="unassigned DNA"
          /db_xref="taxon:10116"
ORIGIN

Qy 290 CACAAACCCAGCGAGCGAGCAATTGAACACTGCACACGGCCCATCTGCCCAGAGAGCTGTGACCAC 349
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8 CACAAACCCAGCGAGCGAGCAATTGAACACTGCACACGGCCCATCTGCCCAGAGAGCTGTGACCAC 67

Qy 350 CACTTCCGCTACTAGCTA 367
        ||||||||||||||| ||
Db   68 CACTTCCGCTACTATCTA 85

RESULT 8
RATIGFB
LOCUS   RATIGFB   1500 bp   mRNA   linear   ROD 27-APR-1993
DEFINITION   Rat IGF binding protein-1 (riGFBP-1) mRNA, comlete cds.
ACCESSION   M58634
VERSION   M58634.1 GI:204732
KEYWORDS   .
SOURCE   IGF binding protein-1.
  ORGANISM   Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE   1 (bases 1 to 1500)
  AUTHORS   Mohn,K.L., Melby,A.E., Tewari,D.S., Laz,T.M. and Taub,R.
  TITLE   The gene encoding rat insulinlike growth factor-binding protein 1
             is rapidly and highly induced in regenerating liver
  JOURNAL   Mol. Cell. Biol. 11 (3), 1393-1401 (1991)
  MEDLINE   91141487
   PUBMED   1705004
COMMENT   Original source text: Rat, cDNA to mRNA.
FEATURES   Location/Qualifiers
  source   1..1500
             /organism="Rattus norvegicus"
             /mol_type="mRNA"
             /strain="Fisher"
             /db_xref="taxon:10116"
             /tissue_type="regenerating liver"
             /dev_stage="adult"
  CDS   160..978
             /codon_start=1
             /product="IGF binding protein-1"
             /protein_id="AAA41380.1"
             /db_xref="GI:204733"
             /translation="MPEFITVVSWPFLILLSFQVRVVAGAPQPWHCAPCTAERLELCP
             PVPASCPEISRPAGCGCPTCALPLGAACGVATARCAQGLSCRALPGERPLHALTRG
             QGACVLEPAAPATSSLSGSGQHEEAKAAVASEDELAESPEMTEEQILLDSFHLMAPSRED
             QPILWNAISYSSMAREITDLKKWKEPCQRELYVLERLAAAQQKAGDEIYKFYLPN
             CNKNGFYHSKQCETSLDGEAGLCWCVYPWSGKKIPGSLETRGDPNCHQYFNVQN"
  mat_peptide 235..975
             /product="IGF binding protein-1"
ORIGIN

Qy 290 CACAAACCCAGCGAGCGAGCAATTGAACACTGCACACGGCCCATCTGCCCAGAGAGCTGTGACCAC 349
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CACAAACCCAGCGAGCGAGCAATTGAACACTGCACACGGCCCATCTGCCCAGAGAGCTGTGACCAC 67

Qy 350 CACTTCCGCTACTAGCTA 367
        ||||||||||||||| ||
Db   68 CACTTCCGCTACTATCTA 85

Query Match   77; Conservative   0; Mismatches   1; Indels   0; Gaps   0;
Best Local Similarity   98.7%; Pred. No. 2.3e-11;
        Score 76.4; DB 6; Length 1500;

RESULT 9
AX163782
LOCUS   AX163782   5001 bp   DNA   linear   PAT 22-JUN-2001
DEFINITION   Sequence 46 from Patent WO0138579.
ACCESSION   AX163782
VERSION   AX163782.1 GI:14544878
KEYWORDS   .
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM   Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE   1
  AUTHORS   Gould-Rothberg,B.E., Dipippo,V.A., Ramseh,T.M. and Gerwein,R.W.
  TITLE   Method of identifying toxic agents using nsaid-induced differential
             gene expression in liver
  JOURNAL   Patent: WO 0138579-A 46 31-MAY-2001;
             Curagen Corporation (US)
FEATURES   Location/Qualifiers
  source   1..5001
             /organism="Rattus norvegicus"
             /mol_type="unassigned DNA"
             /db_xref="taxon:10116"
ORIGIN

Qy 290 CACAAACCCAGCGAGCGAGCAATTGAACACTGCACACGGCCCATCTGCCCAGAGAGCTGTGACCAC 349
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CACAAACCCAGCGAGCGAGCAATTGAACACTGCACACGGCCCATCTGCCCAGAGAGCTGTGACCAC 60

Qy 350 CACTTCCGCTACTAGCTA 367
        ||||||||||||||| ||
Db   61 CACTTCCGCTACTATCTA 78

Query Match   77; Conservative   0; Mismatches   1; Indels   0; Gaps   0;
Best Local Similarity   98.7%; Pred. No. 2.4e-11;
        Score 76.4; DB 6; Length 5001;

RESULT 10
RATIGFBA
LOCUS   RATIGFBA   5001 bp   DNA   linear   ROD 30-NOV-1995
DEFINITION   Rat insulin-like growth factor binding protein-1 (IGFBP-1) gene,
             complete cds.
ACCESSION   L22979
VERSION   L22979.1 GI:1098472
KEYWORDS   insulin-like growth factor binding protein-1.
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM   Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE   1 (bases 1 to 5001)
  AUTHORS   Lacson,R., Oehler,D., Yang,E., Goswami,R. and Unterman,T.
  TITLE   Dideoxy sequencing and structural analysis of the rat insulin-like
             growth factor binding protein-1 gene
  JOURNAL   Biochim. Biophys. Acta 1218 (1), 95-98 (1994)
  MEDLINE   94250701
   PUBMED   7514892
COMMENT   On Nov 30, 1995 this sequence version replaced gi:385167.
             Original source text: Rattus norvegicus (strain Sprague-Dawley)
             DNA.
FEATURES   Location/Qualifiers
  source   1..5001
             /organism="Rattus norvegicus"
             /mol_type="genomic DNA"
             /strain="Sprague-Dawley"

Query Match   77; Conservative   0; Mismatches   1; Indels   0; Gaps   0;
Best Local Similarity   98.7%; Pred. No. 2.3e-11;
        Score 76.4; DB 10; Length 1500;

Qy 290 CACAAACCCAGCGAGCGAGCAATTGAACACTGCACACGGCCCATCTGCCCAGAGAGCTGTGACCAC 349
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
                    /db_xref="taxon:10116"
                    /sex="male"
                    /tissue_type="liver"
                    /dev_stage="adult"
                    /tissue_lib="Clontech EMBL-3 SP6/T7"
     exon           /number=1
                    60..525
     5'UTR          60..152
     gene           153..4221
                    /gene="IGFBP-1"
     CDS            join(153..525,1850..2034,2874..3002,4090..4221)
                    /gene="IGFBP-1"
                    /codon_start=1
                    /product="insulin-like growth factor binding protein"
                    /protein_id="AAA82581.1"
                    /db_xref="GI:1098473"
                    /translation="MPEFLTVVSWPFILLLSRQVRVVAGAPQPWHCAPCTAERLELCP
                    PVPASCPEISRPAGGCGCCPTCALPLGNACGVATAACAQGLSCRALPGEPRPLHALTRG
                    QGACVLEPAAPATSSLSGSQHEEAKAAVASEDELAASSPEMTEEQLLDSFHLMAPSED
                    OPILWNAISTYSSMRAREITDLKKWKEPCQRELYKVLERLAAAQQKAGDEIYKFYLPN
                    CNKNGFYHSKQCETSLDGEAGLCWCVYPWSGKKIPGSLETRGDPNCHQYFNVQN"
     intron         526..1849
                    /gene="IGFBP-1"
                    /number=1
     exon           1850..2034
                    /gene="IGFBP-1"
                    /number=2
     intron         2035..2873
                    /gene="IGFBP-1"
                    /number=2
     exon           2874..3002
                    /gene="IGFBP-1"
                    /number=3
     intron         3003..4089
                    /gene="IGFBP-1"
                    /number=3
     exon           4090..4743
                    /gene="IGFBP-1"
                    /number=4
     3'UTR          4222..4743
ORIGIN

Query Match             20.5%;  Score 76.4;  DB 10;  Length 5001;
Best Local Similarity   98.7%;  Pred. No. 2.4e-11;
Matches    77;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy    290 CACAAACCCAGCGAGGCATTGAACACTGCACACGGCCATCTGCCCAGAGAGCTGTGACCAC 349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CACAAACCCAGCGAGGCATTGAACACTGCACACGGCCATCTGCCCAGAGAGCTGTGACCAC 60

Qy    350 CACTTCCGCCTACTAGCTA 367
          |||||||||| ||||| ||
Db     61 CACTTCCGCTACTATCTA 78
```

```
RESULT 11
AY095345
LOCUS       AY095345                3886 bp    DNA     linear   PRI 30-DEC-2002
DEFINITION  Papio anubis insulin-like growth factor binding protein-1 (IGFBP-1)
            gene, partial cds.
ACCESSION   AY095345
VERSION     AY095345.1  GI:20853764
KEYWORDS    .
SOURCE      Papio anubis (olive baboon)
  ORGANISM  Papio anubis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Papio.
REFERENCE   1  (bases 1 to 3886)
  AUTHORS   Kim,J.J.; Taylor,H.S., Akbas,G.E., Foucher,I., Trembleau,A.,
            Jaffe,R.C., Fazleabas,A.T. and Unterman,T.G.
  TITLE     Regulation of insulin-like growth factor binding protein-1 promoter
            activity by FKHR and HOXA10 in primate endometrial cells
  JOURNAL   Biol. Reprod. 68 (1), 24-30 (2003)
   PUBMED   12493691
REFERENCE   2  (bases 1 to 3886)
  AUTHORS   Kim,J.J., Jaffe,R.C. and Fazleabas,A.T.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-APR-2002) Department of Obstetrics and Gynecology,
            University of Illinois at Chicago, 820 S. Wood, Chicago, IL 60612,
            USA
FEATURES             Location/Qualifiers
     source          1..3886
                     /organism="Papio anubis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9555"
     gene            <3655..>3886
                     /gene="IGFBP-1"
     mRNA            <3655..>3886
                     /gene="IGFBP-1"
     CDS             3655..>3886
                     /gene="IGFBP-1"
                     /codon_start=1
                     /product="insulin-like growth factor binding protein-1"
                     /protein_id="AAM32273.1"
                     /db_xref="GI:20853765"
                     /translation="MSEVPVARVWLVLLLTVQVGVTRSAPWQCAPCSAEKLALCPPV
                     PASCSEVTRSAGGCGCPMCALPLGAACGVATAR"
ORIGIN

Query Match             17.2%;  Score 64;  DB 9;  Length 3886;
Best Local Similarity   60.6%;  Pred. No. 1e-07;
Matches   149;  Conservative   0;  Mismatches  80;  Indels  17;  Gaps   2;

Qy    120 TCCAGGAACCACGGGAGTGCCCCCGTGCCGCCCATGTCACAAGCAAAACAAACTTATTTGA 179
          |||   ||||  |||  || |||  ||||   ||||||||||| |||||||||||||||||
Db   3331 TCCTCCCACCTGCCGTTTGCGTAGGCCTTGGGCGGCCACTRGCGAAAACAAACTTATTTGA 3390

Qy    180 ACACCGGGGATCCTAGCACGCTGCCCTGACAAATCATTAACCC----------GTGCTGCCG 229
          ||||  |   ||||||  |||| |||| ||||||||||| |||          | |||||||
Db   3391 ACACTTCAGCTCCTAGCGTGCGCGCGCTGCCAAATCATTAACCTCCTGGTGCAAGTGGCGCGG 3450

Qy    230 AGCCAGCCCTTCATAAGGCCCTGGGTATGGCCAGCAGCGATGGTCCACTGCCCGCGCCGAGA 289
          |   ||||| ||||||||||  |  | |   |||||||| | ||| || ||||  | |||||
Db   3451 CCTGTGCCCTTTATAAGGCCGCGCGCTGTGTCCAGCAGCATCGGCGCCACCGCCCATC 3510

Qy    290 CACAAACCCAGCGAGGCATTGAACACTGCACACGGGCCATCTGCCCAGAGAGCTGTGACCAC 349
          ||   | |||||||| |||||||| |||| || |  |||||| |||||||||| ||| ||||
Db   3511 CAGCAAGC----------ATCTCCGCGCGCGCGCCCACCCTPCCCAGAGAGGCACTGGCCAC 3563

Qy    350 CACTTC 355
          | |||
Db   3564 CGCTCC 3569
```

```
RESULT 12
AX409747
LOCUS       AX409747                6128 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 2394 from Patent WO0229103.
ACCESSION   AX409747
VERSION     AX409747.1  GI:21442452
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6128)
  AUTHORS   Alvarez,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
  TITLE     Gene expression profiles in liver cancer
  JOURNAL   Patent: WO 0229103-A 2394 11-APR-2002;
            GENE LOGIC INC (US)
FEATURES             Location/Qualifiers
     source          1..6128
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
```

ORIGIN
/note="EMBL/GenBank Accession No. M74587"

Query Match   17.2%; Score 63.8; DB 6; Length 6128;
Best Local Similarity 63.2%; Pred. No. 1.2e-07;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

```
Qy  120 TCCAGGAACCACGGGAGTGCCCCGTGCGCCCATGTCACAAGCAAAACAAAACTTATTTTGA 179
            |||  ||| |||||||||||  |||
Db  438 TCCTCCCACCAGCGGTTTGCGTAGGGCCCTTGGGTGCACTAGCAAAAACAAACTTATTTTGA 497

Qy  180 ACACGCGGGATCCTAGCACGCGCTGCCCTTGACAATCATTAACCC------------GTGCTGCCG 229
Db  498 ACACTCAGCTCCTAGCGTGCGCGCTGCCAATCATTAACCTCCTTGGTGCAAGTGGCGCGG 557

Qy  230 AGCCAGCGCCTTCATAAGGCCCTGGGGTATGGCCAGCCAGCATGGTTCCACTGCCCGCCGAGA 289
Db  558 CCTGTGCCCTTTATAAGGTGCGCGCTGTGTCCAGCGCATCGGCGCCACCGCCATCC---- 613

Qy  290 CACAAACCCACGCGAGCATTGAACACTGC-ACACGGCCCATCTGCCCAGAGAGCTGTGACCA 348
Db  614 ----CATCCAGCGAGCATCTGCCGCCGCCCGCGCCGCCACCCTCCCAGAGAGCACTGGCCA 669

Qy  349 CCACTTC 355
Db  670 CCGCTCC 676
```

RESULT 13
HUMIGFBP1A

LOCUS   HUMIGFBP1A   6128 bp   DNA   linear   PRI 08-NOV-1994
DEFINITION Human insulin-like growth factor binding protein (hIGFBP1) gene,
            complete cds.
ACCESSION   M74587
VERSION   M74587.1 GI:184811
KEYWORDS   .
SOURCE   insulin-like growth factor binding protein.
  ORGANISM Homo sapiens
            Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 6128)
  AUTHORS   Ehrenborg,E., Larsson,C., Stern,I., Janson,M., Powell,D.R. and
            Luthman,H.
  TITLE   Contiguous localization of the genes encoding human insulin-like
            growth factor binding proteins 1 (IGBP1) and 3 (IGBP3) on
            chromosome 7
  JOURNAL   Genomics 12 (3), 497-502 (1992)
  MEDLINE   92217971
  PUBMED   1373120
COMMENT   Original source text: Homo sapiens (tissue library: lambda
            Charon4A) liver DNA.
FEATURES   Location/Qualifiers
     source   1..6128
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="7p13-p12"
                     /tissue_type="liver"
                     /tissue_lib="lambda Charon4A"
     CAAT_signal   525..529
                     /gene="IGFBP1"
     TATA_signal   569..573
                     /gene="IGFBP1"
     gene   join(597..1110,2657..2826,4041..4169,5069..5769)
                     /gene="IGFBP1"
                     /note="G00-120-075"
     mRNA   join(597..1110,2657..2826,4041..4169,5069..5769)
                     /gene="IGFBP1"
                     /product="insulin-like growth factor binding protein 1"
                     /note="G00-120-075"
     exon   597..1110
                     /gene="IGFBP1"

     CDS   join(762..1110,2657..2826,4041..4169,5069..5200)
                     /gene="IGFBP1"
                     /codon_start=1
                     /product="insulin-like growth factor binding protein 1"
                     /protein_id="AAA52784.1"
                     /db_xref="GI:184812"
                     /db_xref="GDB:G00-120-075"
                     /translation="MSEVPVARVWLVLLLLTVQVGVTAGAPWQCAPCSAEKLALCPPV
                     SASCSEVTRSAGCGCCPMCALPLGAACGVATFARCARGLSCRALPGEQPLHALTRGQG
                     ACVQESDASAPHAAEAGSPESPESTEITEEELLDNFHLMAPSEEDHSILWDAISTYDG
                     SKALHVTNIKKWKEPCRIELYRVVESLAKAQETSGEEISKFYLPNCNKNGFYHSRQCE
                     TSMDGEAGLCWCVYPWNGKRIPGSEIRGDPNCQMYFNVQN"
     sig_peptide   762..836
                     /gene="IGFBP1"
     mat_peptide   join(837..1110,2657..2826,4041..4169,5069..5197)
                     /gene="IGFBP1"
                     /product="insulin-like growth factor binding protein 1"
                     /note="G00-120-075"
     exon   4041..4169
                     /gene="IGFBP1"
                     /note="G00-120-075"
     intron   2827..4040
                     /gene="IGFBP1"
                     /note="G00-120-075"
     intron   1111..2656
                     /gene="IGFBP1"
                     /note="G00-120-075"
     intron   4170..5068
                     /gene="IGFBP1"
                     /note="G00-120-075"
     variation   4736
                     /gene="IGFBP1"
                     /note="G00-120-075"
     variation   4780
                     /gene="IGFBP1"
                     /note="G00-120-075"
     variation   4820
                     /gene="IGFBP1"
                     /note="G00-120-075"
     exon   5069..5769
                     /gene="IGFBP1"
                     /note="G00-120-075"
     variation   5069..5769
                     /gene="IGFBP1"
                     /note="G00-120-075"
     exon   2657..2826
                     /gene="IGFBP1"
                     /note="G00-120-075"
     intron   1405
                     /gene="IGFBP1"
                     /note="G00-120-075"

ORIGIN

Query Match   17.2%; Score 63.8; DB 9; Length 6128;
Best Local Similarity 63.2%; Pred. No. 1.2e-07;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

```
Qy  120 TCCAGGAACCACGGGAGTGCCCCGTGCGCCCATGTCACAAGCAAAACAAAACTTATTTTGA 179
            |||  ||| |||||||||||  |||
Db  438 TCCTCCCACCAGCGGTTTGCGTAGGGCCCTTGGGTGCACTAGCAAAAACAAACTTATTTTGA 497

Qy  180 ACACGCGGGATCCTAGCACGCGCTGCCCTTGACAATCATTAACCC------------GTGCTGCCG 229
Db  498 ACACTCAGCTCCTAGCGTGCGCGCTGCCAATCATTAACCTCCTTGGTGCAAGTGGCGCGG 557

Qy  230 AGCCAGCGCCTTCATAAGGCCCTGGGGTATGGCCAGCCAGCATGGTTCCACTGCCCGCCGAGA 289
Db  558 CCTGTGCCCTTTATAAGGTGCGCGCTGTGTCCAGCGCATCGGCGCCACCGCCATCC---- 613

Qy  290 CACAAACCCACGCGAGCATTGAACACTGC-ACACGGCCCATCTGCCCAGAGAGCTGTGACCA 348
Db  614 ----CATCCAGCGAGCATCTGCCGCCGCCCGCGCCGCCACCCTCCCAGAGAGCACTGGCCA 669

Qy  349 CCACTTC 355
```

RESULT 14

```
LOCUS       G19994                  6128 bp   DNA   linear   STS 28-SEP-1998
DEFINITION  G19994 swS8752 Eric D. Green Homo sapiens STS genomic, sequence tagged
            site.
ACCESSION   G19994
VERSION     G19994.1   GI:1254693
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 6128)
  AUTHORS   Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F.,
            Weintraub,L.A., Mohr-Tidwell,R.M., Peluso,D.C., Fulton,R.S.,
            Leckie,M.P. and Green,E.D.
  TITLE     A collection of 1814 human chromosome 7-specific STSs
  JOURNAL   Genome Res. 7 (1), 59-64 (1997)
  MEDLINE   97189344
   PUBMED   9037602
REFERENCE   2 (bases 1 to 6128)
  AUTHORS   Green,E.D.
  TITLE     Human chromosome 7 STSs (1997)
  JOURNAL   Unpublished (1997)
COMMENT     Synonyms: IGFBP1
            GDB: GDB:3754042
            GDB_DSEG: IGFBP1
            Contact: Eric D. Green
            Genome Technology Branch
            National Human Genome Research Institute/NIH
            49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
            Tel: 301024020201
            Fax: 301024735
            Email: egreen@nhgri.nih.gov
            Primer A: TATCAGACAGACAGGTAGG
            Primer B: TGAGGAACCCAAGATCCAAG
            STS size: 340
            PCR Profile:
              Presoak:          0 degrees C for 0.00 minute(s)
              Denaturation:    92 degrees C for 1.00 minute(s)
              Annealing:       60 degrees C for 2.00 minute(s)
              Polymerization:  72 degrees C for 2.00 minute(s)
              PCR Cycles:      35
              Thermal Cycler:  PerkinElmer TC
            Protocol:
              Template:         30-100 ng
              Primer:           each 1 uM
              dNTPs:            each 200 uM
              Taq Polymerase:   0.05 units/ul
              Total Vol:        5 ul
            Buffer:
              MgCl2:            1.5 mM
              KCl:              50 mM
              Tris-HCl:         10 mM
              pH:               8.3
            This STS was developed from sequence determined by another
            investigator. See GenBank record: M74587  For additional
            information about the NHGRI chromosome 7 mapping project, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7.  Also see Genomics
            11:548-64 (1991) [MUID:92128937].
FEATURES             Location/Qualifiers
     source          1..6128
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="7"
                     /clone_lib="Eric D. Green"
```

RESULT 15

```
Query Match          17.2%;  Score 63.8;  DB 11;  Length 6128;
Best Local Similarity 63.2%;  Pred. No. 1.2e-07;
Matches  156;  Conservative   0;  Mismatches  72;  Indels  19;  Gaps  3;

ORIGIN

Qy  120 TCCAGGAACCACGGGAGTGCCCCGTGCGCCCCATGTCACAAGCAAAACAAACTTATTTGA 179
         |||||  |||||| ||||| |||||||||    |||||||| |||||||||||||||||||
Db  438 TCCTCCCACCACGCGGTTGCGGTAGGGCCTTGGGTGCACTAGCAAAACAAACTTATTTGA 497

Qy  180 ACACCGGGGAATCCTAGCACGCTGCCCTGACAATCATTAACCC------GTGCTGCCG 229
         |||| | |  |||||||| | || |||  ||||||||||||         || || ||
Db  498 ACACTCAGTCCTAGCGTGGGGCGGCTGCCAATCATTAACCTCCTGGTGCAAGTGCGCGG 557

Qy  230 AGCCAGCCCTTCATTAAGGCCCCTGGGTATGCCACGCCAGCATGGTCCACTGCCCCGCCGAGA 289
         | | ||||| ||||||||| || |||| ||  || |||||||| | ||| || || ||||
Db  558 CCTGTGCCCTTTATTAAGGTCCGCCGCGTGTCGTCCAGCGAGCATCGGCGCCACCGCCATCC 613

Qy  290 CACAAACCCAGCGAGCCATTGAACACTGC-ACACGGCCCATCTGCCCAGGAGGAGCGTGTGACCA 348
             |||||||||| |||||| |  | ||| | |||||||| || || ||||||||||   |||| |||
Db  614 ----CATCCAGCGGAGCATCTGCCGCCGCCGCGCCCGGCCACCCTCCCAGGAGGAGCACTGGCCA 669

Qy  349 CCACTTC 355
         ||| ||
Db  670 CCCGCTCC 676
```

```
LOCUS       HUMIGFBP1               6480 bp   DNA   linear   PRI 08-NOV-1994
DEFINITION  Human insulin-like growth factor binding protein-1 (IGFBP1) gene,
            complete cds.
ACCESSION   M59316
VERSION     M59316.1   GI:184809
KEYWORDS    insulin-like growth factor binding protein 1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 6480)
  AUTHORS   Cubbage,M.L., Suwanichkul,A. and Powell,D.R.
  TITLE     Structure of the human chromosomal gene for the 25 kilodalton
            insulin-like growth factor binding protein
  JOURNAL   Mol. Endocrinol. 3 (5), 846-851 (1989)
  MEDLINE   89330502
   PUBMED   2474129
REFERENCE   2 (bases 1 to 6480)
  AUTHORS   Suwanichkul,A., Cubbage,M.L. and Powell,D.R.
  TITLE     The promoter of the human gene for insulin-like growth factor
            binding protein-1. Basal promoter activity in HEP G2 cells depends
            upon liver factor B1
  JOURNAL   J. Biol. Chem. 265 (34), 21185-21193 (1990)
  MEDLINE   91065933
   PUBMED   1701175
COMMENT     Original source text: Human leukocyte DNA.
FEATURES             Location/Qualifiers
     source          1..6480
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="7p13-p12"
                     /cell_type="leukocyte"
     gene            join(1206..1719,3258..3427,4643..4771,5662..6362)
                     /gene="IGFBP1"
     mRNA            join(1206..1719,3258..3427,4643..4771,5662..6362)
                     /gene="IGFBP1"
```

Middle-column feature/match annotations:

```
              1..6128
              /gene="IGFBP1"
     STS      4155..4494
              /gene="IGFBP1"
     primer_bind  4155..4174
              /gene="IGFBP1"
     primer_bind  complement(4475..4494)
              /gene="IGFBP1"
```

```
Db    670 CCCGCTCC 676
```

```
     exon            1206..1719
                     /gene="IGFBP1"
                     /note="G00-120-075; putative"
                     /number=1
     CDS             join(1371..1445,1446..1719,3258..3427,4643..4771,
                     5662..5793)
                     /gene="IGFBP1"
                     /note="putative"
                     /note="G00-120-075"
                     /codon_start=1
                     /product="insulin-like growth factor binding protein-1"
                     /protein_id="AAA52783.1"
                     /db_xref="GI:184810"
                     /db_xref="GDB:G00-120-075"
                     /translation="MSEVPVARVWLVLLLLTVQVGVTAGAPWQCAPCSAEKLALCPPV
                     SASCSEVTRSAGCGCCPMCALPLGAACGVATARCARGLSCRALPGEQQPLHALTRGQG
                     ACVQESDASAPHAAEAGSPESPESTEITEEELLDNFHLMAPSEEDHSILWDAISTYDG
                     SKALHVTNIKKWKEPCRIELYRVVESLAKAQETSGEEISKFYLPNCNKNGFYHSRQCE
                     TSMDGEAGLCWCVYPWNGKRIPGSPEIRGDPNCQIYFNVQN"
     sig_peptide     1371..1445
                     /gene="IGFBP1"
                     /note="G00-120-075"
     mat_peptide     join(1446..1719,3258..3427,4643..4771,5662..5790)
                     /gene="IGFBP1"
                     /evidence=experimental
                     /product="insulin-like growth factor binding protein-1"
                     /note="G00-120-075; putative"
     exon            3258..3427
                     /gene="IGFBP1"
                     /note="G00-120-075; putative"
                     /number=2
     exon            4643..4771
                     /gene="IGFBP1"
                     /note="G00-120-075; putative"
                     /number=3
     exon            5662..6362
                     /gene="IGFBP1"
                     /note="G00-120-075; putative"
                     /number=4
ORIGIN          Chromosome 7.

Query Match           17.2%;  Score  63.8;  DB  9;  Length 6480;
Best Local Similarity 63.2%;  Pred. No. 1.2e-07;
Matches  156;  Conservative  0;  Mismatches  72;  Indels  19;  Gaps  3;

Qy   120 TCCAGGAACCACGGGGAGCTGCCCGTGCGCCCCATGTCACAAGGCAAAACAAAACTTATTTTGA 179
          ||||  ||||| |||||||| |||||   |||||| |||| |||| | || |||||||||||||
Db  1047 TCCTCCCACCAGCGGGTTTGCCGTAGGCGCCTTGGGTGCACTAGCAAAACAAACTTATTTTTGA 1106

Qy   180 ACACGGGGATCCTAGCACGCTGCCCTGACAATCATTAACCC---------GTGCTGCCG 229
          |||| | | |||||||| |  | |||| |  ||| |||| |          |||||||||
Db  1107 ACACTCAGCTCCTAGCCGTGCGGCGGCTGCCAATCATTAACCTTCCTGGTTGCAAGTGGGCCGG 1166

Qy   230 AGCCAGCCCCTTCATAAGGGCCCTGGGGTATGGCCAGCCAGCCATGGTCCACTGCCCGCCGGAGA 289
          |||||||||| ||||| |  | | |||  || |||| ||| || ||||| |||||| ||||||
Db  1167 CCTGTGCCCTTTATAAGGTGCGCGGCTGTGTCCGAGCACCGCCATCC---- 1222

Qy   290 CACAAACCCAGCGGAGCATTGAACACTGC-ACACGGCGCCATCTGCCCAGAGAGCTGTGACCA 348
          |||||||| |||| || ||| | |||| | |||| |||| || ||| ||||||||| | ||||
Db  1223 ----CATCCAGCGAGCATCTCCGCGCGCCGCCGCGCGCCGCCACCCTCCCAGAGAGCACTGGCCA 1278

Qy   349 CCACTTC 355
          || | ||
Db  1279 CCGGCTCC 1285

Search completed: June  7, 2004, 13:58:20
Job time : 2647 secs
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: June 7, 2004, 08:46:36 ; Search time 339 Seconds
(without alignments)
4661.739 Million cell updates/sec

Title: US-09-972-916B-5
Perfect score: 372
Sequence: 1 tacactggggccagagtcc..........ttccgctactagctagccgc 372
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 3373863 seqs, 212409041 residues
Total number of hits satisfying chosen parameters: 6747726
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2010s:*
5: geneseqn2001bg:*
6: geneseqn2001bg:*
7: geneseqn2002aa:*
8: geneseqn2003bg:*
9: geneseqn2003cg:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	372	100.0	372	6	ABX15375	Abx15375 Rat insul
2	321	86.3	321	6	ABX15374	Abx15374 Rat insul
3	223.6	60.1	270	6	ABX15376	Abx15376 Rat insul
4	219.4	59.0	219	6	ABX15373	Abx15373 Rat insul
5	219	58.9	219	6	ABX15372	Abx15372 Rat insul
c 6	154	41.4	423	6	ABX15376	Abx15376 Rat insul
7	76.4	20.5	1500	6	ABK63701	Abk63701 Rat seque
8	76.4	20.5	1500	7	ABT41911	Abt41911 Toxicity
9	76.4	20.5	1500	6	ABT41911	Abt41911 Toxicity-
10	76.4	20.5	1500	9	ADB58201	Adb58201 Toxicity
11	76.4	20.5	1501	4	ADB52710	Adb52710 Primary r
12	63.8	17.2	6128	7	AAH22429	Aah22429 Rat insul
13	63.8	17.2	6128	4	ABN95896	Abn95896 Gene #239
14	63.8	17.2	6134	4	ABV75371	Abv75371 Human IGF
15	55.2	14.8	194	2	AAH57489	Aah57489 Human liv
16	55.2	14.8	13011	2	AAT43002	Aat43002 Rat type
17	55.2	14.8	13011	7	AAT96631	Aat96631 cDNA enco
18	51	13.7	51	7	ABT42448	Abt42448 Toxicity
19	51	13.7	51	6	ABX15380	Abx15380 Rat liver
20	51	13.7	270	6	ABX15371	Abx15371 Rat gluco
21	48	12.9	270	6	ABX15373	Abx15373 Rat insul
22	38.6	10.4	4861	6	ABX15379	Abx15379 Rat insul
23	37.8	10.2	7061	6	AAS20000	Aas20000 DNA encod
					ABL32997	Abl32997 Human imm
c 24	37.8	10.2	7061	6	ABL70248	Abl70248 Chemical1
c 25	37.8	10.2	7061	6	AAS61189	Aas61189 Human gen
c 26	37.8	10.1	1254	5	AAS77024	Aas77024 DNA encod
c 27	37.4	10.1	12850	6	ADB36326	Adb36326 Human fac
c 28	36.8	9.9	761	6	ABL90266	Abl90266 Human pol
c 29	36.6	9.8	1678	7	ADA52441	Ada52441 Human cod
c 30	36.6	9.8	3054	6	ABS70481	Abs70481 Human bon
c 31	35.8	9.6	11009	6	ABQ72907	Abq72907 Mouse lam
c 32	35.8	9.6	11009	6	AAI70816	Aai70816 Mouse lam
c 33	35.6	9.6	490	8	ACH34858	Ach34858 Human end
c 34	35	9.4	3895	6	ABN87725	Abn87725 Human pro
c 35	34	9.1	10537	6	AAK69582	Aak69582 Human car
c 36	34	9.1	10537	6	AAK69582	Aak69582 Human car
c 37	34	9.1	10537	6	ADE46463	Ade46463 Human car
c 38	34	9.1	10543	6	AAS35770	Aas35770 Human car
c 39	34	9.1	10543	6	AAK69583	Aak69583 Human imm
c 40	34	9.1	10543	9	ADE46464	Ade46464 Human car
c 41	34	9.1	26277	8	AAK70428	Aak70428 Human imm
c 42	33.4	9.0	24873	8	ADA02594	Ada02594 Human BMI
c 43	33.4	9.0	24873	9	ADB72332	Adb72332 Human BMI
c 44	33.4	9.0	28000	9	ADD00956	Add00956 Human Jag
c 45	33.2	8.9	650	7	ABZ52705	Abz52705 Aspergill

ALIGNMENTS

RESULT 1
ABX15375

ID   ABX15375 standard; DNA; 372 BP.
XX
AC   ABX15375;
XX
DT   17-APR-2003 (first entry)
XX
DE   Rat insulin regulator construct DNA #3.
XX
KW   Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW   insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW   insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW   fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW   glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW   hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW   intracellular hormone receptor; insulin regulator construct; anabolic.
OS   Rattus norvegicus.
OS   Synthetic.
XX
PN   US2002107198-A1.
XX
PD   08-AUG-2002.
XX
PF   10-OCT-2001; 2001US-00972916.
XX
PR   11-OCT-2000; 2000US-0239113P.
XX
PT   (THUL/) THULE P M.
PT   liver cells.
XX
PA   (THUL/) THULE P M.
XX
PI   Thule PM;
XX
DR   WPI; 2002-674190/72.
XX
PP   New insulin regulator cassette, useful e.g. for treating diabetes, provides specific, glucose-inducible transgenic expression of insulin in liver cells.
XX
PS   Claim 9; Page 14; 37pp; English.
XX
CC   The invention relates to an insulin regulator construct comprising a glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene promoter and an insulin-sensitive element (ISE) of an insulin-like growth factor binding protein-1 (IGFBP-1) basal promoter. The construct is used to treat or prevent diabetic complications, to regulate insulin

CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
SQ Sequence 372 BP; 87 A; 128 C; 102 G; 55 T; 0 U; 0 Other;

Query Match   100.0%; Score 372; DB 6; Length 372;
Best Local Similarity 100.0%; Pred. No. 4.9e-103; Mismatches 0;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TACACTGGGGGGGCCAGAGTCCAGGAACCACGGGAGTGCCCCGTGCGCCCCATGTACACTGGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TACACTGGGGGGGCCAGAGTCCAGGAACCACGGGAGTGCCCCGTGCGCCCCATGTACACTGGG 60

Qy   61 GGCCAGAGTCCAGGAACCACGGGAGTGCCCCGTGCGCCCCATGTACACTGGGGGGCCAGAGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GGCCAGAGTCCAGGAACCACGGGAGTGCCCCGTGCGCCCCATGTACACTGGGGGGCCAGAGT 120

Qy   121 CCAGGAACCACGGGAGGAGTGCCCCGTGCCCCATGTCAAGAGCAAAAACAAAACTTATTTTTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCAGGAACCACGGGAGGAGTGCCCCGTGCCCCATGTCAAGAGCAAAAACAAAACTTATTTTGAA 180

Qy   181 CACGGGGGATCCTTAGCACGGCTGCCCCTGACAATCATTAACCCGTGCTCGCGAGCCCAGCCCTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CACGGGGGATCCTTAGCACGGCTGCCCCTGACAATCATTAACCCGTGCTCGCGAGCCCAGCCCTT 240

Qy   241 CATAAGGCCCTGGGTATGGCCAGCCACCAGCAGCCATGGTCCACTGCCCCGCCGAGACACAAACCCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CATAAGGCCCTGGGTATGGCCAGCCACCAGCAGCCATGGTCCACTGCCCCGCCGAGACACAAACCCAG 300

Qy   301 CGAGCAATTGAACACTGCACACGGCCATCTGCCCAGAGAGCTGTGACACCACACACTTCCGCTA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CGAGCAATTGAACACTGCACACGGCCATCTGCCCAGAGAGCTGTGACACCACACACTTCCGCTA 360

Qy   361 CTAGCTAGCCGC 372
          ||||||||||||
Db   361 CTAGCTAGCCGC 372

XX
OS Rattus norvegicus.
OS Synthetic.
PN US20020107198-A1.
XX

RESULT 2
ABX15374
ID ABX15374 standard; DNA; 321 BP.
XX
AC ABX15374;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #2.
XX
KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX

──

PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
PA (THUL/) THULE P M.
XX
PI Thule PM;
XX
DR WPI; 2002-674190/72.
XX
PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Claim 9; Page 14; 37pp; English.
XX
CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis;
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
SQ Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;

Query Match   86.3%; Score 321; DB 6; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.5e-87; Mismatches 0;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   52 TACACTGGGGGGGCCAGAGTCCAGGAACCACGGGAGTGCCCCGTGCGCCCCATGTACACTGGG 111
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TACACTGGGGGGGCCAGAGTCCAGGAACCACGGGAGTGCCCCGTGCGCCCCATGTACACTGGG 60

Qy   112 GGCCAGAGTCCAGGAACCACGGGAGTGCCCCGTGCCCCGTGCGCCCCATGTCACAGCAAAAGCAAAAACAAAACT 171
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GGCCAGAGTCCAGGAACCACGGGAGTGCCCCGTGCCCCGTGCGCCCCATGTCACAGCAAAAGCAAAAACAAAACT 120

Qy   172 TATTTTGAACACGGGGGATCCTTAGCACGGCTGCCCCTGACAATCATTAACCCGTGCTCGCGAG 231
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TATTTTGAACACGGGGGATCCTTAGCACGGCTGCCCCTGACAATCATTAACCCGTGCTCGCGAG 180

Qy   232 CCAGCCCTTCATTAAGGCCCCTGGGTATGGCCAGCCACCAGCAGCCATGGTCCACTGCCCCGCCGAGACA 291
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCAGCCCTTCATTAAGGCCCCTGGGTATGGCCAGCCACCAGCAGCCATGGTCCACTGCCCCGCCGAGACA 240

Qy   292 CAAACCCAGCGAGCAATTGAACACTGCACACGGCCATCTGCCCAGAGAGCTGTGACACCACCA 351
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAAACCCAGCGAGCAATTGAACACTGCACACGGCCATCTGCCCAGAGAGCTGTGACACCACCA 300

Qy   352 CTTCCGCTACTAGCTAGCCGC 372
          |||||||||||||||||||||
Db   301 CTTCCGCTACTAGCTAGCCGC 321

Db
XX

RESULT 3
ABX15376
ID ABX15376 standard; DNA; 423 BP.
XX

AC   ABX15376;
XX
DT   17-APR-2003   (first entry)
XX
DE   Rat insulin regulator construct DNA #4.
XX
KW   Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW   insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW   insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW   fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW   glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW   hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW   intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS   Rattus norvegicus.
OS   Synthetic.
XX
PN   US2002107198-A1.
XX
PD   08-AUG-2002.
XX
PF   10-OCT-2001; 2001US-00972916.
XX
PR   11-OCT-2000; 2000US-0239113P.
XX
PA   (THUL/) THULE P M.
XX
PI   Thule PM;
XX
DR   WPI; 2002-674190/72.
XX
PT   New insulin regulator cassette, useful e.g. for treating diabetes,
PT   provides specific, glucose-inducible transgenic expression of insulin in
PT   liver cells.
XX
PS   Claim 9; Page 14; 37pp; English.
XX
CC   The invention relates to an insulin regulator construct comprising a
CC   glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC   promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC   factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC   to treat or prevent diabetic complications, to regulate insulin
CC   production, to modulate hyperglycaemia (without severe hypoglycaemia),
CC   increase fat catabolism and to retard protein catabolism. The construct
CC   provides stimulation of insulin expression by glucose (but not other
CC   carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC   combination of these effects can maintain nearly euglycaemic conditions
CC   in diabetics during short-term fasting, large carbohydrate loads or when
CC   fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC   thus inhibiting the long-term complications of diabetes. The properties
CC   of the construct are essentially specific for hepatocytes and well-
CC   differentiated hepatoma lines and insulin expression in these cells may
CC   have effects additional to those provided by secreted insulin, e.g.
CC   inhibition of cellular protein degradation, and inhibition, or
CC   stimulation of other intracellular hormone receptors. This sequence
CC   represents a rat insulin regulator construct of the invention
XX
SQ   Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;

Query Match   60.1%; Score 223.6; DB 6; Length 423;
Best Local Similarity   76.2%;
Matches 297; Conservative 0; Mismatches 74; Indels 19; Gaps 1;

```
Qy   2 ACACTGGGGCGCCCAGAGTTCCAGGAACCACAGGGAGTGCCCCGTGGGCGCCCAGTCGGGG 61
         ||||||||||||||||| |||||||||||||||||| ||||||||
Db  34 ACTCTGGGCCCCCAGTGTACATGGGCGCACGGGGCCACTCCCGTGGTTCCTGGACTCTGGCC 93

Qy  62 GCCAGAGTCCAGGAACCACGGGAGTGCCCCGTGGGCGGGGGGCCCAGTCACTGTACACTGC 121
         |||||  |||||| || |||  | ||||||||||||  ||||||  |||
Db  94 CCCAGTGTACATGGGCGCACGGGGCCACTCCCGTGGTTCCTGGACTCTGGCC CCAGTC 153

Qy 122 CAGGAACCACGGGAGTGCCCCGTGCGCCCCATG------------------TCACAAGCA 162
         |||||||||                                        |||||||||
```

RESULT 4
ABX15373
ID   ABX15373 standard; DNA; 270 BP.
XX
AC   ABX15373;
XX
DT   17-APR-2003   (first entry)
XX
DE   Rat insulin regulator construct DNA #1.
XX
KW   Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW   insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW   insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW   fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW   glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW   hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW   intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS   Rattus norvegicus.
OS   Synthetic.
XX
PN   US2002107198-A1.
XX
PD   08-AUG-2002.
XX
PF   10-OCT-2001; 2001US-00972916.
XX
PR   11-OCT-2000; 2000US-0239113P.
XX
PA   (THUL/) THULE P M.
XX
PI   Thule PM;
XX
DR   WPI; 2002-674190/72.
XX
PT   New insulin regulator cassette, useful e.g. for treating diabetes,
PT   provides specific, glucose-inducible transgenic expression of insulin in
PT   liver cells.
XX
PS   Claim 9; Page 13-14; 37pp; English.
XX
CC   The invention relates to an insulin regulator construct comprising a
CC   glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC   promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC   factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC   to treat or prevent diabetic complications, to regulate insulin
CC   production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC   increase fat catabolism and to retard protein catabolism. The construct
CC   provides stimulation of insulin expression by glucose (but not other
CC   carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC   combination of these effects can maintain nearly euglycaemic conditions
CC   in diabetics during short-term fasting, large carbohydrate loads or when
CC   fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC   thus inhibiting the long-term complications of diabetes. The properties

```
Db 154 CATGGGGCGCACGGGGCGCACTCCGTCGGTTCTCGGACTCTGGCCCCCAGTGTATCACAAGCA 213
Qy 163 AAACAAAACTTATTTGAACACGGGGGATCCTAGCACGCGTGCCCTGACAATCATTAACCCGT 222
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 214 AAACAAAACTTATTTGAACACGGGGGATCCTAGCACGCGTGCCCTGACAATCATTAACCCGT 273

Qy 223 GCTGCGGACGCCAGCCCCTTCATTAAGCCCCTGGGTATGGGCCAGCACACCATGGTTCCACTGCCC 282
         ||||| |||||||||||||||||| ||||||||||||||  ||| ||  |||||||||||||||
Db 274 GCTGCCGAGCCCAGCCCCTTCATTAAGGCCCTGGGTATGGCCAGCCACCCAGCCATGGTTCCACTGCCC 333

Qy 283 GCCGAGACACACAAACCCAGCGCATTGAACACTGCACACCGGCCCCAGAGAGCTG 342
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 334 GCCGAGACACACAAACCCAGCGCATTGAACACTGCACACCGGCCCCAGAGAGCTG 393

Qy 343 TGACCACCACTTCCGCTACTAGCTAGCCGC 372
         ||||||||||||||||||||||||||||||
Db 394 TGACCACCACTTCCGCTACTAGCTAGCCGC 423
```

CC   of the construct are essentially specific for hepatocytes and well-
CC   differentiated hepatoma lines and insulin expression in these cells may
CC   have effects additional to those provided by secreted insulin, e.g.
CC   inhibition of cellular protein degradation, and inhibition, or
CC   stimulation of other intracellular hormone receptors. This sequence
CC   represents a rat insulin regulator construct of the invention
XX
SQ   Sequence 270 BP; 63 A; 95 C; 65 G; 47 T; 0 U; 0 Other;

Query Match   59.0%; Score 219.4; DB 6; Length 270;
Best Local Similarity   99.5%; Pred. No. 1e-56;
Matches 220; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   152 TGTCACAAGCAAAACAAAACTTATTTTGAACACGGGGGATCCTAGCACCGTTGCCCTGACAAT 211
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   50 TATCACAAGCAAAACAAAACTTATTTTGAACACGGGGGATCCTAGCACCGTTGCCCTGACAAT 109

Qy   212 CATTAACCCGTTGCCGGAGCCAGCCCTTCATAAGGCCCTGGGTATGGCCAGCCAGCCATG 271
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   110 CATTAACCCGTTGCCGGAGCCAGCCCTTCATAAGGCCCTGGGTATGGCCAGCCAGCCATG 169

Qy   272 GTCCACTGCCCCGCCGAGACACAAAACCCAGCGAGCAATTGAACACTGCACACGGCCCATCTGC 331
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   170 GTCCACTGCCCCGCCGAGACACAAAACCCAGCGAGCAATTGAACACTGCACACGGCCCATCTGC 229

Qy   332 CCAGAGAGCTGTGACCACCACTTCCGCTACTTAGCTAGCCGC 372
          ||||||||||||||||||||||||||||||||||||||||||
Db   230 CCAGAGAGCTGTGACCACCACTTCCGCTACTTAGCTAGCCGC 270

RESULT 5
ABX15372
ID   ABX15372 standard; DNA; 219 BP.
XX
AC   ABX15372;
XX
DT   17-APR-2003 (first entry)
XX
DE   Rat insulin-sensitive element (ISE) DNA.
XX
KW   Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW   insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW   insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW   fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW   glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW   hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW   intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS   Rattus norvegicus.
XX
DE   Rat insulin-sensitive element (ISE) DNA.
XX
PN   US2002107198-A1.
XX
DR   08-AUG-2002.
XX
PF   10-OCT-2001; 2001US-00972916.
XX
PR   11-OCT-2000; 2000US-0239113P.
XX
PA   (THUL/) THULE P M.
XX
PI   Thule PM;
XX
DR   WPI; 2002-674190/72.

Claim 8; Page 13; 37pp; English.

New insulin regulator cassette, useful e.g. for treating diabetes,
provides specific, glucose-inducible transgenic expression of insulin in
liver cells.

The invention relates to an insulin regulator construct comprising a
glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
promoter and an insulin-sensitive element (ISE) of an insulin-like growth

---

CC   factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC   to treat or prevent diabetic complications, to regulate insulin
CC   production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC   increase fat catabolism and to retard protein catabolism. The construct
CC   provides stimulation of insulin expression by glucose (but not other
CC   carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC   combination of these effects can maintain nearly euglycaemic conditions
CC   in diabetics during short-term fasting, large carbohydrate loads or when
CC   fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC   thus inhibiting the long-term complications of diabetes. The properties
CC   of the construct are essentially specific for hepatocytes and well-
CC   differentiated hepatoma lines and insulin expression in these cells may
CC   have effects additional to those provided by secreted insulin, e.g.
CC   inhibition of cellular protein degradation, and inhibition, or
CC   stimulation of other intracellular hormone receptors. This sequence
CC   represents a rat insulin-sensitive element of the invention
XX
SQ   Sequence 219 BP; 57 A; 77 C; 48 G; 37 T; 0 U; 0 Other;

Query Match   58.9%; Score 219; DB 6; Length 219;
Best Local Similarity   100.0%; Pred. No. 1.3e-56;
Matches 219; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   154 TCACAAGCAAAACAAAACTTATTTTGAACACGGGGGATCCTAGCACCGTTGCCCTGACAATCA 213
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TCACAAGCAAAACAAAACTTATTTTGAACACGGGGGATCCTAGCACCGTTGCCCTGACAATCA 60

Qy   214 TTAACCCGTTGCCGGAGCCAGCCCTTCATAAGGCCCTGGGTATGGCCAGCCAGCCATGGT 273
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTAACCCGTTGCCGGAGCCAGCCCTTCATAAGGCCCTGGGTATGGCCAGCCAGCCATGGT 120

Qy   274 CCACTGCCCCGCCGAGACACAAAACCCAGCGAGCAATTGAACACTGCACACGGCCCATCTGCC 333
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCACTGCCCCGCCGAGACACAAAACCCAGCGAGCAATTGAACACTGCACACGGCCCATCTGCC 180

Qy   334 AGAGAGCTGTGACCACCACTTCCGCTACTTAGCTAGCCGC 372
          ||||||||||||||||||||||||||||||||||||||||
Db   181 AGAGAGCTGTGACCACCACTTCCGCTACTTAGCTAGCCGC 219

RESULT 6
ABX15376/c
ID   ABX15376 standard; DNA; 423 BP.
XX
AC   ABX15376;
XX
DT   17-APR-2003 (first entry)
XX
DE   Rat insulin regulator construct DNA #4.
XX
KW   Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW   insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW   insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW   fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW   glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW   hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW   intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS   Rattus norvegicus.
OS   Synthetic.
XX
PN   US2002107198-A1.
XX
PD   08-AUG-2002.
XX
PF   10-OCT-2001; 2001US-00972916.
XX
PR   11-OCT-2000; 2000US-0239113P.
XX
PA   (THUL/) THULE P M.
XX
PI   Thule PM;
XX

```
DR    WPI; 2002-674190/72.
XX
PT    New insulin regulator cassette, useful e.g. for treating diabetes,
PT    provides specific, glucose-inducible transgenic expression of insulin in
PT    liver cells.
XX
PS    Claim 9; Page 14; 37pp; English.
XX
CC    The invention relates to an insulin regulator construct comprising a
CC    glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC    promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC    factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC    to treat or prevent diabetic complications, to regulate insulin
CC    production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC    increase fat catabolism and to retard protein catabolism. The construct
CC    provides stimulation of insulin expression by glucose (but not other
CC    carbohydrates) and glucocorticoids and inhibition of insulin expression by
CC    combination of these effects can maintain nearly euglycaemic conditions
CC    in diabetics during short-term fasting, large carbohydrate loads or when
CC    fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC    thus inhibiting the long-term complications of diabetes. The properties
CC    of the construct are essentially specific for hepatocytes and well-
CC    differentiated hepatoma lines and insulin expression in these cells may
CC    have effects additional to those provided by secreted insulin, e.g.
CC    inhibition of cellular protein degradation, and inhibition, or
CC    stimulation of other intracellular hormone receptors. This sequence
CC    represents a rat insulin regulator construct of the invention
XX
SQ    Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;

Query Match             41.4%;  Score 154;  DB 6;  Length 423;
Best Local Similarity  100.0%;  Pred. No. 8.9e-37;
Matches 154;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 TACACTGGGGGCGGGCCAGAGTCCAGGAACCACGGGAGGAGTGCCCCGTCGCCCCAGGAACTGGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    204 TACACTGGGGGCGGGCCAGAGTCCAGGAACCACGGGAGGAGTGCCCCGTGGGCCATCAGTACACTGGG 145

Qy     61 GGGCCAGAGTCCAGGAACCACGGGAGGAGTGCCCCGTGGGGGCCCAGAGT 120
          ||||||||||||||||||||||||||||||||||||||||||||
Db    144 GGGCCAGAGTCCAGGAACCACGGGAGGAGTGCCCCCATGTACACTGGGGGGGGCCAGAGT 85

Qy    121 CCAGGAACCACGGGAGGAGTGCCCCGTGCGCCCCATGT 154
          ||||||||||||||||||||||||||||||||||||||
Db     84 CCAGGAACCACGGGAGGAGTGCCCCGTGCGCCCCATGT 51


RESULT 7
ABK63701
ID   ABK63701 standard; cDNA; 1500 BP.
XX
AC   ABK63701;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Rat sequence differentially expressed in response to a hepatotoxin #1608.
XX
KW   Rat; 88; hepatotoxin; expressed sequence tag; EST; drug screening;
KW   differential expression; centrilobular necrosis; steatosis.
XX
OS   Rattus norvegicus.
XX
PN   WO200210453-A2.
XX
PD   07-FEB-2002.
XX
PF   30-JUL-2001; 2001WO-US023872.
XX
PR   31-JUL-2000; 2000US-0222040P.
PR   02-NOV-2000; 2000US-0244880P.
PR   11-MAY-2001; 2001US-0290029P.
PR   15-MAY-2001; 2001US-0290645P.
PR   22-MAY-2001; 2001US-0292336P.
PR   06-JUN-2001; 2001US-0295798P.
PR   13-JUN-2001; 2001US-0297457P.
PR   19-JUN-2001; 2001US-0298884P.
PR   09-JUL-2001; 2001US-0303459P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX
DR   WPI; 2002-241625/29.
XX
PT   Predicting toxic effects of compounds or the progression of these toxic
PT   effects by determining the changes in gene expression in tissues or cells
PT   exposed to the toxin and comparing these to gene expression in unexposed
PT   tissues or cells.
XX
PS   Claim 1; SEQ ID NO 1608; 239pp; English.
XX
CC   The invention relates to methods for predicting toxic effects of
CC   compounds or the progression of these toxic effects by determining the
CC   global changes in gene expression in tissues or cells exposed to the
CC   toxin and comparing these to gene expression in unexposed tissues or
CC   cells. Also included are methods of predicting at least one toxic effect
CC   of a compound or progression of a toxic effect, preferably the
CC   hepatotoxicity of a compound, comprising detecting the level of
CC   expression in a tissue or cell sample exposed to the compound of two or
CC   more genes listed in the specification, where differential expression of
CC   the genes is indicative of at least one toxic effect or progression. The
CC   method can also be used to identify an agent which modulates the toxic
CC   response and predict cellular pathways that a compound modulates in a
CC   cell. The methods utilise a set of at least two probes (on a solid
CC   support in kit form), where each of the probes comprises a sequence that
CC   specifically hybridises to a gene listed in the specification, a computer
CC   system comprising a database containing information identifying the
CC   expression level in a tissue or cell sample exposed to a hepatotoxin of a
CC   set of genes comprising at least two genes listed in the specification,
CC   and a user interface to view the information used to present information
CC   identifying the expression level in a tissue or cell of at least one gene
CC   listed in the specification. The method is useful for elucidating global
CC   changes in gene expression and for identifying toxicity markers in
CC   tissues or cell exposed to a known toxin. The genes may be used as
CC   toxicity markers in drug screening and toxicity assays. The genes and
CC   gene expression information may be used as diagnostic markers for the
CC   prediction or identification of the physiological state of tissue or cell
CC   sample that has been exposed to a compound or agent. Hepatotoxicity is
CC   characterised by centrilobular necrosis and steatosis. The present
CC   sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC   which is differentially expressed in response to a hepatotoxic agent
XX
SQ   Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match            20.5%;  Score 76.4;  DB 6;  Length 1500;
Best Local Similarity  98.7%;  Pred. No. 5e-13;
Matches 77;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

Qy    290 CACAAAACCCAGCGAGGCATTGAACAACTGCACACGGCCATCTGCCCAGGAGAGCTGTGACCAC 349
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8 CACAAAACCCAGCGAGGCATTGAACAACTGCACACGGCCATCTGCCCAGGAGAGCTGTGACCAC 67

Qy    350 CACTTCCGCTACTAGCTA 367
          |||||||||||| |||
Db     68 CACTTCCGCTACTATCTA 85


RESULT 8
ABT41911
ID   ABT41911 standard; DNA; 1500 BP.
XX
AC   ABT41911;
XX
DT   26-JUN-2003  (first entry)
XX
DE   Toxicity modelling related rat gene SEQ ID No 1613.
```

```
XX
KW   Toxic effect; gene expression profile; renal toxicity; toxicity marker;
KW   database; drug screening; toxicity assay; rat; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO20295000-A2.
XX
PD   28-NOV-2002.
XX
PF   22-MAY-2002; 2002WO-US016173.
XX
PR   22-MAY-2001; 2001US-0292335P.
PR   13-JUN-2001; 2001US-0297523P.
PR   19-JUN-2001; 2001US-0298825P.
PR   10-JUL-2001; 2001US-0303806P.
PR   10-JUL-2001; 2001US-0303807P.
PR   10-JUL-2001; 2001US-0303808P.
PR   28-AUG-2001; 2001US-0315047P.
PR   27-SEP-2001; 2001US-0324928P.
PR   22-OCT-2001; 2001US-0330462P.
PR   01-NOV-2001; 2001US-0330867P.
PR   21-NOV-2001; 2001US-0331805P.
PR   06-DEC-2001; 2001US-0336144P.
PR   19-DEC-2001; 2001US-0340873P.
PR   21-FEB-2002; 2002US-0357842P.
PR   21-FEB-2002; 2002US-0357843P.
PR   21-FEB-2002; 2002US-0357844P.
PR   15-MAR-2002; 2002US-0364134P.
PR   08-APR-2002; 2002US-0370144P.
PR   08-APR-2002; 2002US-0370206P.
PR   08-APR-2002; 2002US-0370247P.
PR   17-APR-2002; 2002US-0372794P.
PR   21-APR-2002; 2002US-0371679P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX
DR   WPI; 2003-148464/14.
XX
PT   Predicting at least one toxic effect of a compound, useful for toxicity
PT   modeling, comprises preparing a gene expression profile of a tissue or
PT   cell sample exposed to the compound, and comparing the gene expression
PT   profile to a database.
XX
PS   Example 4; Page; 446pp; English.
XX
CC   The invention relates to a novel method of predicting at least one toxic
CC   effect of a compound. The method comprises a gene expression profile of a
CC   tissue or cell sample exposed to the compound, and comparing the gene
CC   expression profile to a database comprising at least part of the data or
CC   information given in the specification. The methods are useful for
CC   predicting at least one toxic effect of a compound, predicting the
CC   progression of a toxic effect of a compound, predicting the renal
CC   toxicity of a compound, or identifying toxicity markers in tissues or
CC   cells exposed to a compound. The genes are useful as toxicity
CC   markers in drug screening and toxicity assays, in monitoring disease or
CC   physiological states, or disease progression. This polynucleotide
CC   represents a rat DNA sequence relating to the toxic effect database
CC   described in the specification. NOTE: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from the World Intellectual Property
CC   Organization
XX
SQ   Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
XX

Query Match          20.5%; Score 76.4; DB 7; Length 1500;
Best Local Similarity  98.7%; Pred. No. 5e-13;
Matches  77; Conservative  0; Mismatches  1; Indels  0; Gaps  0;

Qy  290 CACAAACCCAGGCGAGCATTGAACACTGCACACGGCCATTCGCCCAGAGAGCTGTGACCAC 349
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db
```

```
Db     8 CACAAACCCAGGCGAGCATTGAACACTGCACACGGCCATTCTGCCCAGAGAGCTGTGACCAC 67

Qy   350 CACTTCCGCTACTTAGCTA 367
         |||||||||| |||
Db    68 CACTTCCGCTACTATCTA 85

RESULT 9
ADB58201
ID   ADB58201 standard; DNA; 1500 BP.
XX
AC   ADB58201;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Toxicity-related gene, SEQ ID 3227.
XX
KW   Toxic; toxin; gene expression profile; hepatotoxicity; liver;
KW   drug screening; toxicity assay; ds.
XX
OS   Unidentified.
XX
PN   WO2003064624-A2.
XX
PD   07-AUG-2003.
XX
PF   31-JAN-2003; 2003WO-US003194.
XX
PR   31-JAN-2002; 2002US-00060087.
PR   15-MAR-2002; 2002US-03640459.
PR   15-MAR-2002; 2002US-0364055P.
PR   30-DEC-2002; 2002US-0436643P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX
DR   WPI; 2003-689530/65.
XX
PT   Predicting a toxic effect of a compound, useful in identifying toxicity
PT   markers in liver tissues or cells for drug screening and toxicity assays,
PT   comprises preparing gene expression profile of tissue or cells exposed to
PT   the compound.
XX
PS   Claim 1; SEQ ID NO 3227; 1156pp; English.
XX
CC   The present invention relates to a method for predicting a toxic effect
CC   of a compound. The method comprises preparing a gene expression profile
CC   of a tissue or cell sample exposed to the compound, and comparing the
CC   gene expression profile to a database comprising SEQ ID 1-4925, where
CC   differential expression of the gene indicates at least one toxic effect.
CC   The method is useful for predicting at least one toxic effect of a
CC   compound, predicting hepatotoxicity or the progression of a toxic effect
CC   of a compound, identifying an agent that modulates the onset or
CC   progression of a toxic response, predicting the cellular pathways that a
CC   compound modulates in a cell, and identifying an agent that modulates at
CC   least one activity of a protein. The method and compositions of the
CC   present invention using a database of genes having liver toxin-induced
CC   differential expression, are useful in identifying liver toxin-induced
CC   liver tissues or cells for drug screening and toxicity assays. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
XX

Query Match          20.5%; Score 76.4; DB 9; Length 1500;
Best Local Similarity  98.7%; Pred. No. 5e-13;
Matches  77; Conservative  0; Mismatches  1; Indels  0; Gaps  0;

Qy  290 CACAAACCCAGGCGAGCATTGAACACTGCACACGGCCATTCGCCCAGAGAGCTGTGACCAC 349
        ||||||||||||||||||||||||||||| |||||||||| |||||||||||||||||||||
Db    8 CACAAACCCAGGCGAGCATTGAACACTCCACACGGCCATCTGCCCAGAGAGCTGTGACCAC 67
```

Qy 350 CACTTCCGCTACTAGCTA 367
        |||||||||||||||| |||
Db   68 CACTTCCGCTACTATCTA 85

RESULT 10
ID   ADB52710 standard; DNA; 1500 BP.
XX
AC   ADB52710;
XX
DT   04-DEC-2003 (first entry)
XX
DE   Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3252.
XX
KW   toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW   toxicity marker; toxicity progression; drug screening;
KW   primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO2003065993-A2.
XX
PD   14-AUG-2003.
XX
PF   04-FEB-2003; 2003WO-US003482.
XX
PR   04-FEB-2002; 2002US-0353171P.
PR   13-MAR-2002; 2002US-0363534P.
PR   08-APR-2002; 2002US-0370248P.
PR   10-APR-2002; 2002US-0371134P.
PR   10-APR-2002; 2002US-0371135P.
PR   10-APR-2002; 2002US-0371150P.
PR   11-APR-2002; 2002US-0371413P.
PR   19-APR-2002; 2002US-0373601P.
PR   19-APR-2002; 2002US-0373602P.
PR   22-APR-2002; 2002US-0374139P.
PR   08-MAY-2002; 2002US-0378370P.
PR   09-MAY-2002; 2002US-0378652P.
PR   09-MAY-2002; 2002US-0378653P.
PR   09-MAY-2002; 2002US-0378665P.
PR   09-MAY-2002; 2002US-0378665P.
PR   09-JUL-2002; 2002US-0394230P.
PR   09-JUL-2002; 2002US-0394253P.
PR   04-SEP-2002; 2002US-0407680P.
PR   28-JAN-2003; 2003US-0442900P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI   Elashoff M;
XX
DR   WPI; 2003-731472/69.
XX
PT   Determining if a compound induces a toxic effect on a tissue or cell, for
PT   identifying hepatotoxic compounds, comprises comparing a gene expression
PT   profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT   mean values.
XX
PS   Claim 44; SEQ ID NO 3252; 874pp; English.
XX
CC   The present invention describes a method for determining whether a
CC   compound induces a toxic effect on a tissue or cell. The method comprises
CC   preparing a gene expression profile of a tissue or cell sample exposed to
CC   the compound, and comparing the gene expression profile to a database
CC   comprising data or information on the Tox mean and non-Tox mean value.
CC   The method is useful for predicting or identifying at least one toxic
CC   effect, particularly hepatotoxicity, of a test or unknown compound. The
CC   genes listed in the specification are useful as diagnostic or toxicity
CC   markers for the prediction or identification of the physiological state
CC   of a tissue or cell sample that has been exposed to a compound, or to
CC   identify or predict the toxic effects of a compound or an agent. These
CC   may also be used as markers for monitoring toxicity progression or for
CC   drug screening. The present sequence represents a primary rat hepatocyte
CC   toxicity modelling related gene sequence from the present invention.
XX
SQ   Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match   20.5%; Score 76.4; DB 9; Length 1500;
Best Local Similarity   98.7%;
Matches   77; Conservative   0; Mismatches   1; Indels   0; Gaps   0;
                                  Pred. No. 5e-13;

Qy 290 CACAAACCCAGGACGCAGCCATTGAAACACTGCACACGGCCATTCTGCCCCAGAGAGAGCTGTGACCAC 349
        ||||||| || || |||||| |||| |||||||||||||||| | ||||||||| ||||||||||||||
Db   8 CACAAACCCACCGAGCGAGCATTGAACACTGCACACGGCCATCTCTCCCAGAGAGCTGTGACCAC 67

Qy 350 CACTTCCGCTACTAGCTA 367
        |||||||||||||| |||
Db   68 CACTTCCGCTACTATCTA 85

RESULT 11
ID   AAH22429 standard; DNA; 5001 BP.
XX
AC   AAH22429;
XX
DT   22-AUG-2001 (first entry)
XX
DE   Rat insulin-like growth factor binding protein nucleotide sequence.
XX
KW   Identification; toxic; hepatotoxic; differential gene expression; NSAID;
KW   non-steroidal antinflammatory drug; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO200138579-A2.
XX
PD   31-MAY-2001.
XX
PF   21-NOV-2000; 2000WO-US032049.
XX
PR   22-NOV-1999; 99US-0166923P.
PR   18-FEB-2000; 2000US-0183531P.
PR   20-NOV-2000; 2000US-00717321.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Gould-Rochberg BE, Dipippo VA, Ramseh TM, Gerwein RW;
XX
DR   WPI; 2001-355948/37.
XX
PT   Screening hepatotoxic agent comprises contacting test cell population
PT   expressing RISKMARKER or INJURYMARKER with agent, comparing expression
PT   with reference population and identifying difference in expression
PT   levels.
XX
PS   Disclosure; Page 22-24; 76pp; English.
XX
CC   The present invention describes a method of screening a test agent for
CC   hepatotoxicity. The method comprises: (a) providing a test cell
CC   population comprising a cell capable of expressing one or more nucleic
CC   acid sequences selected from the group consisting of RISKMARKER 1-8 and
CC   INJURYMARKER 1-10; (b) contacting the test cell population with a test
CC   agent; (c) measuring expression of one or more of the nucleic acid
CC   sequences in the test cell population; (d) comparing the expression of
CC   the nucleic acid sequence in the test cell population to the expression
CC   of the nucleic acid sequence in an reference cell population comprising
CC   at least one cell whose exposure status to a hepatotoxic agent is known;
CC   and (e) identifying a difference in expression levels of the RISKMARKER
CC   or INJURYMARKER sequences, if present, in the test cell population and
CC   reference cell population. The method is useful for identifying a
CC   hepatotoxic agent. The present sequence is given in the exemplification
CC   of the present invention
XX
SQ   Sequence 5001 BP; 1225 A; 1204 C; 1261 G; 1311 T; 0 U; 0 Other;

```
Query Match    20.5%; Score 76.4; DB 4; Length 5001;
Best Local Similarity  98.7%;
Matches  77; Conservative  0; Mismatches  1; Indels  0; Gaps  0;

Qy  290 CACAAACCCACGCGAGCANTTGAACACTTGCACACGGCCATTCTGCCCAGAGAGCTGTGACCAC 349
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 CACAAACCCACGCGAGCANTTGAACACTTGCACACGGCCATTCTGCCCAGAGAGCTGTGACCAC 60

Qy  350 CACTTCCGCCTACTAGCTA 367
         |||||||||||| |||
Db   61 CACTTCCGCTACTATCTA 78
```

RESULT 12
ABN95896

```
ID   ABN95896 standard; DNA; 6128 BP.
XX
AC   ABN95896;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Gene #2394 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
XX
PN   WO20029103-A2.
XX
PI   Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
PD   11-APR-2002.
XX
PF   02-OCT-2000; 2000US-0237054P.
PR   02-OCT-2000; 2000US-0237054P.
XX
PT   (GENE-) GENE LOGIC INC.
XX
PS   Claim 1; SEQ ID NO 2394; 298pp; English.
XX
DR   WPI; 2002-426119/45.
XX
CC   The invention relates to a novel method for diagnosing and detecting the
CC   progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC   tumour in a patient, and differentiating metastatic liver cancer from
CC   hepatocellular carcinoma in a patient, involving detecting the level of
CC   expression of two or more genes represented in ABN93503-ABN97455 in a
CC   tissue sample. The method of the invention has hepatotropic, and
CC   cytostatic activity. The method is useful for diagnosing and detecting
CC   the progression of liver cancer, hepatocellular carcinoma and metastatic
CC   liver carcinoma in a patient. The method is useful for identifying
CC   expression profiles which serve as useful diagnostic markers as well as
CC   markers that can be used to monitor disease states, disease progression,
CC   drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;
```

```
Query Match    17.2%; Score 63.8; DB 6; Length 6128;
Best Local Similarity  63.2%;
Matches  156; Conservative  0; Mismatches  72; Indels  19; Gaps  3;

Qy  120 TCCAGGAACCACGGGAGGAGTGCCCCGTGCGCCCATGTCACAAGCAAAGCAAAACAAACTTATTTGA 179
```

```
Qy  438 TCCTCCCACCAGCGGGTTTGCGTAGGGCCTTGGGTGCCACTAGCAAAAACAAAACTTATTTGA 497
         |||||  |||  |||       |||          |||  |||        || ||||||||||||
Db  180 ACACGGGGGATCCTTAGCACGCGTGCCCTAGCACAAATCATTAACCC----------GTGCTGCG 229

Qy  498 ...                                                               557
Db  230 AGCCAGCCCTTCATTAAGCGCCCTGGGTATGGCCAGCCAGCAGGCATGGTTCCACTGCCCGCCGAGA 289

Qy  558 ...                                                               613
Db  290 CCTGGTGCCCTTTATTAAGGTGCGGGCTGTGTCCACGGCATCGGCCACCGCGCCCATCC---- 348

Qy  614 ----CATCCAGCGAGCAACTCGCCGCCGCCGCCGCCGCGGCGCACCCTCCCAGAGAGCACTGGGCA 669
Db  349 ...                                                                 355

Qy  349 CCACTTC 355
         ||| |||
Db  670 CCGCTCC 676
```

RESULT 13
ABV75371

```
ID   ABV75371 standard; DNA; 6128 BP.
XX
AC   ABV75371;
XX
DT   07-MAR-2003  (first entry)
XX
DE   Human IGFBP-1 gene sequence.
XX
KW   Insulin-like growth factor binding protein; IGFBP; cytostatic; liver;
KW   cancer; human; IGFBP-1; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT
FT   CDS             762..5200
FT                   /product= "IGFBP-1"
FT                   /note= "insulin-like growth factor binding protein;
FT                   contains introns"
FT   exon            762..1110
FT                   /number= 1
FT                   /*tag= b
FT   intron          1111..2656
FT                   /number= 1
FT                   /*tag= c
FT   exon            2657..2826
FT                   /number= 2
FT                   /*tag= d
FT   intron          2827..4040
FT                   /number= 2
FT                   /*tag= e
FT   exon            4041..4169
FT                   /number= 3
FT                   /*tag= f
FT   intron          4170..5068
FT                   /number= 3
FT                   /*tag= g
FT   exon            5069..5197
FT                   /number= 4
FT                   /*tag= h
XX
PN   WO20029580-A1.
XX
PD   14-NOV-2002.
XX
PF   03-MAY-2002; 2002WO-AU000558.
PR   03-MAY-2001; 2001US-0288441P.
XX
PA   (NACA-) NAT CANCER CENT SINGAPORE PTE LTD.
PA   (SIGE-) SINGAPORE GEN HOSPITAL PTE LTD.
```

PA   (ARGA/) ARGAET V P.
XX
PI   Huynh TH, Chow PKH, Soo KC;
XX
DR   WPI; 2003-103522/09.
DR   P-PSDB; ABB82757.
XX
PT   Detecting the presence or diagnosing the risk of a liver cancer in a
PT   patient comprises detecting aberrant expression of a gene encoding an
PT   insulin-like growth factor binding protein.
XX
PS   Example; Page 104-108; 142pp; English.
XX
CC   The invention relates to detecting the presence or diagnosing the risk of
CC   a liver cancer in a patient. The method involves detecting in a
CC   biological sample obtained from the patient aberrant expression of a gene
CC   encoding an insulin-like growth factor binding protein (IGFBP). The
CC   method is useful for detecting the presence or diagnosing the risk of a
CC   liver cancer or for screening agents in a patient. The agent is useful
CC   for the manufacture of a medicament for treating and/or preventing liver
CC   cancer. The present sequence represents a human IGFBP-1 polypeptide
CC   encoding genomic DNA (GenBank Accession No. M74587).
XX
SQ   Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;

Query Match   17.2%; Score 63.8; DB 7; Length 6128;
Best Local Similarity   63.2%; Pred. No. 5.2e-09;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

Qy   120 TCCAGGAACCACGGGAGGGAGTGCCCCGTGCGCCCATGTCACAAGCAAAACAAACTTATTTGA 179
              |||| ||| ||||| ||| |||| | || |||| ||||||||||||||||||||
Db   438 TCCTCCCACCACGCGGTTTGCGTAGGGGTGCACTAGCAAAACAAACTTATTTGA 497

Qy   180 ACACGGGGGATCCTAGCACCCTGCCCTGACAATCATTATTAACCC----------GTGCTGCCG 229
              ||||||| || ||| | | |||||||||||||||||||| |||||||||
Db   498 ACACTCAGCTCCTAGCGTGCGGCCGCTGCCAATCATTAACCTCCTCGGTGTGCAAGTGCGCG 557

Qy   230 AGCCAGCGCCCTTCATTAAGGCCCTGGGTATGGCGCCAGCCAGCCATGGTCCACTGCCCGGCCGGAGA 289
              ||||||||||||||||||||||||||||||||||||||||| |||||||||| ||||||| |||||||
Db   558 CCTGTGCCCTTTATTAAGGTCGCGCCTGTGTCCACGGACGCATCGGCCACCGCCACCCATCC---- 613

Qy   290 CACAAACCCAGCGGAGCATTGAAACACTGC-ACACGGCCCATCTGCCCAGAGAGGCTGTGACCA 348
              |||||||||||||||||||||||| |||| |||||||||| ||| ||||| ||||||||||||
Db   614 ----CATCCAGCGAGCATCTGCCCGCGCGCGGGCGCGGAGCAGAGCACTGGGCCA 669

Qy   349 CCACTTC 355
              || |||
Db   670 CCGGTCC 676

XX
ID   AAH57489 standard; cDNA; 6134 BP.
XX
AC   AAH57489;
XX
DT   10-SEP-2001 (first entry)
XX
DE   Human liver cell specific cDNA sequence SEQ ID NO:329.
XX
KW   Human; tissue specific; diagnosis; brain; heart; skeletal muscle; lung;
KW   liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW   metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW   neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
OS   Homo sapiens.
XX
PN   WO200132927-A2.
XX
PD   10-MAY-2001.
XX
PF   02-NOV-2000; 2000WO-US030396.
XX

PR   04-NOV-1999; 99US-01635080P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Sornasse T, Seilhamer JJ, Watson GA;
XX
WPI; 2001-291057/30.
XX
SQ   Sequence 6134 BP; 1505 A; 1498 C; 1507 G; 1624 T; 0 U; 0 Other;

Query Match   17.2%; Score 63.8; DB 4; Length 6134;
Best Local Similarity   63.2%; Pred. No. 5.3e-09;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

Qy   120 TCCAGGAACCACGGGAGGGAGTGCCCCGTGCGCCCATGTCACAAGCAAAACAAACTTATTTGA 179
              |||| ||| ||||| ||| |||| | || |||| ||||||||||||||||||||
Db   438 TCCTCCCACCACGCGGTTTGCGTAGGGGTGCACTAGCAAAACAAACTTATTTGA 497

Qy   180 ACACGGGGGATCCTAGCACCCTGCCCTGACAATCATTATTAACCC----------GTGCTGCCG 229
              ||||||| || ||| | | |||||||||||||||||||| |||||||||
Db   498 ACACTCAGCTCCTAGCGTGCGGCCGCTGCCAATCATTAACCTCCTCGGTGTGCAAGTGCGCG 557

Qy   230 AGCCAGCGCCCTTCATTAAGGCCCTGGGTATGGCGCCAGCCAGCCATGGTCCACTGCCCGGCCGGAGA 289
              ||||||||||||||||||||||||||||||||||||||||| |||||||||| ||||||| |||||||
Db   558 CCTGTGCCCTTTATTAAGGTCGCGCCTGTGTCCACGGACGCATCGGCCACCGCCACCCATCC---- 613

Qy   290 CACAAACCCAGCGGAGCATTGAAACACTGC-ACACGGCCCATCTGCCCAGAGAGGCTGTGACCA 348
              |||||||||||||||||||||||| |||| |||||||||| ||| ||||| ||||||||||||
Db   614 ----CATCCAGCGAGCATCTGCCCGCGCGCGGGCGCGGAGCAGAGCACTGGGCCA 669

Qy   349 CCACTTC 355
              || |||
Db   670 CCGGTCC 676

XX

RESULT 15
ID   AAT43002/c
ID   AAT43002 standard; DNA; 194 BP.
XX
AC   AAT43002;
XX
DT   16-JUL-1997 (first entry)
XX
DE   Rat type L pyruvate kinase promoter region.
XX
KW   Glucose-inducible; rat; pyruvate kinase type L; L-PK; promoter;
KW   hyperglycaemia; diabetes; gene therapy; viral vector; defective virus;
KW   ss.
XX
OS   Ratus sp.
XX

```
FH   Key            Location/Qualifiers
FT   promoter       16..39
FT                  /*tag= a
FT                  /label= L4_element
FT                  /note= "able to interact with MLTF/USF factor"
FT   promoter       40..58
FT                  /*tag= b
FT                  /label= L3_element
FT                  /notes= "able to interact with HNF 4 factor"
XX
FT
PN   WO9632489-A1.
XX
PD   17-OCT-1996.
XX
PF   12-APR-1996;  96WO-FR000560.
XX
PR   14-APR-1995;  95FR-00004558.
XX
PA   (INRM ) INST NAT SANTE & RECH MEDICALE.
XX
PI   Chen R,  Doiron B,  Kahn A;
XX
DR   WPI; 1996-477139/47.
XX
PT   Recombinant defective virus contg. heterologous gene under control of
PT   glucose inducible promoter - and infected mammalian cells, esp. for gene
PT   therapy of diabetes.
XX
PS   Claim 4; Page 21; 41pp; French.
XX
CC   A new defective recombinant virus contains at least one heterologous gene
CC   under control of an expression signal inducible by glucose or its
CC   analogues. Preferably, the expression signal is derived from the 183 bp
CC   region located 5' of the type L pyruvate kinase (L-PK) gene, especially
CC   having the present sequence. Expression of the recombinant virus can be
CC   controlled by glucose levels, but insulin production can be stopped
CC   immediately by administering glucagon. This avoids the risk of
CC   hypoglycaemia caused by excessive insulin production. Mammalian cells
CC   infected by the virus can be implanted into liver, spleen, pancreas or
CC   intestine, to provide insulin secretion at predetermined sites. The virus
CC   is useful in gene therapy to treat and/or prevent diseases associated
CC   with hyperglycaemia, particularly diabetes, but more generally can be
CC   used to provide controlled expression of a wide variety of proteins
XX
SQ   Sequence 194 BP; 50 A; 59 C; 51 G; 34 T; 0 U; 0 Other;
```

Query Match   14.8%; Score 55.2; DB 2; Length 194;
Best Local Similarity   88.2%;
Matches   60; Conservative   0; Mismatches   8; Indels   0; Gaps   0;

```
Qy         35 TGCCCCGTGCGCCCATGTAACACTGGGGGGGGGGGCCAGAGTCCAGGAACCACGGGAGTGCCCCGTG 94
              |||
Db         78 TGCCAACGGAAGCCTTGTACACTGGGGGCGCCAGAGTCCAGGAACCACGGGAGTGCCCCGTG 19

Qy         95 CGCCCATG 102
              ||||||||
Db         18 CGCCCATG 11
```

Search completed: June 7, 2004, 14:04:12
Job time : 341 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 11:17:56 ; Search time 2491 Seconds
                                           (without alignments)
                                           4459.547 Million cell updates/sec

Title: US-09-972-916B-5
Perfect score: 372
Sequence: 1 tacactgggggccagagtcc..........ttccgctactagctagccgc 372
Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 27513289 seqs, 1493109276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Databases :   EST:*
             1: em_estba:*
             2: em_esthum:*
             3: em_estin:*
             4: em_estmu:*
             5: em_estov:*
             6: em_estpl:*
             7: em_estro:*
             8: em_htc:*
             9: gb_est1:*
            10: gb_est2:*
            11: gb_htc:*
            12: gb_est3:*
            13: gb_est4:*
            14: gb_est5:*
            15: em_estfun:*
            16: em_estcom:*
            17: em_gss_hum:*
            18: em_gss_inv:*
            19: em_gss_pln:*
            20: em_gss_vrt:*
            21: em_gss_fun:*
            22: em_gss_mam:*
            23: em_gss_mus:*
            24: em_gss_pro:*
            25: em_gss_rod:*
            26: em_gss_phg:*
            27: em_gss_vrl:*
            28: gb_gss1:*
            29: gb_gss2:*

                              SUMMARIES

Result   %
No.   Score   Query   Match Length DB   ID   Description
----- ----- ------- ----- ------ -- -------- -----------
  1   98.2   26.4   1013   10   BF236974   BF236974 60202660
  2   87.8   23.6   310   13   BY783538   BY783538 BY783538
  3   87.8   23.6   331   13   BY794229   BY794229 BY794229
  4   87.8   23.6   354   13   BY771317   BY771317 BY771317

5	86.8	23.3	401	9	AI785818	AI785818 ui78h05.y
6	86.8	23.3	480	9	AI196314	AI196314 ui71a07.y
7	86.8	23.3	706	9	AI530146	AI530146 ui89f09.y
8	86.8	23.3	765	9	AI098594	AI098594 ue31e07.y
9	86.8	23.3	785	9	AI530313	AI530313 ui91f01.y
10	86.8	23.3	799	9	AI529939	AI529939 ui87c09.y
11	86.8	23.3	848	9	AI790802	AI790802 uk28b10.y
12	86.8	23.3	852	9	AI528304	AI528304 ui95g10.y
13	85.2	22.9	605	9	AI196154	AI196154 ui69d08.y
14	82.2	22.1	380	9	AI785039	AI785039 ui73a06.y
15	73.2	19.7	332	14	W30013	W30013 mc24c07.r1
16	72.4	19.5	615	10	BB660958	BB660958 BB660958
17	72.4	19.5	618	14	CD561711	CD561711 B043SD02-
18	72.4	19.5	632	9	AI892189	AI892189 mj88b01.y
19	67.4	18.1	488	9	AA674302	AA674302 vp96g10.r
20	62.8	16.9	759	9	AA103355	AA103355 mp37d09.r
21	61.8	16.6	269	10	BB604790	BB604790 BB604790
22	61.6	16.6	587	10	AW916227	AW916227 EST347531
23	59	15.9	545	9	AA060360	AA060360 mj67a12.r
24	47.8	12.8	1159	29	CNS015XR	CNS015XR Drosophil
25	47.2	12.7	659	28	AZ847793	AZ847793 2M018D02
26	42.2	11.3	597	28	AZ652514	AZ652514 1M052SH14
27	41.6	11.1	325	10	AW855818	AW855818 RC1-HT025
28	40	10.8	212	10	BE149547	BE149547 RC1-HT025
29	39.8	10.7	1201	13	BX381961	BX381961 BX381961
30	39.4	10.6	500	28	AQ612869	AQ612869 HS_5116_A
31	39.2	10.5	234	10	AW886850	AW886850 RC1-OT008
32	39.2	10.5	289	10	BE066031	BE066031 RC3-BT031
33	39.2	10.5	1101	29	CNS017SY	CNS017SY Drosophil
34	39.2	10.5	476	14	CB062462	CB062462 4012331 B
35	39	10.5	479	14	CB061591	CB061591 4014411 B
36	38.8	10.4	698	13	BU684425	BU684425 UI-CF-EN0
37	38.8	10.4	425	10	AL803306	AL803306 AL803306
38	38.6	10.4	547	13	BQ601057	BQ601057 MI-P-H1-a
39	38.4	10.3	307	10	AW886846	AW886846 RC1-OT008
40	38.4	10.3	1201	13	BX361080	BX361080 BX361080
41	38	10.2	925	29	CNS00091P	CNS00091P Drosophil
42	37.8	10.2	304	9	AW062732	AW062732 CM0-CT010
43	37.8	10.2	1103	12	BM916417	BM916417 AGENCOURT
44	37.6	10.1	624	13	BX858228	BX858228 BX858228
45	37.6	10.1	1040	13	BQ954857	BQ954857 AGENCOURT

                              ALIGNMENTS

RESULT 1
BF236974

FEATURES
  source   Location/Qualifiers
                  1. .1013
                  /organism="Mus musculus"
                  /mol_type="mRNA"
                  /strain="FVB/N"
                  /db_xref="taxon:10090"
                  /clone="IMAGE:4161889"
                  /lab_host="DH10B (T1 phage-resistant)"
                  /clone_lib="NCI_CGAP_Li9"
                  /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
                  Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT;
                  Average insert size 1.9 kb. Constructed by Life
                  Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

RESULT 2
BY783538

Query Match   Best Local Similarity   92.8%;   Score 98.2;   DB 10;   Length 1013;
                                               Pred. No. 5.8e-18;
Matches 103;   Conservative   0;   Mismatches   8;   Indels   0;   Gaps   0;

Qy 257 TGGCCAGCGCAATGGTCCACTCCCCGCCGGAGACAACAAAACCCAGCGAGCCATTGAACACTG 316
            |||||||| ||||||||||||| |||||| |||||||| |||||||||||||||||||||||
Db   9 TGGGCAGCCAGCAATGGTCCACTGCCCGCGGAGACAACACACCCAGCGAGCCATTGAACACTG 68

Qy 317 CACACGGCCATCTGCCCAGAGAGAGGCTGTGACCACCACTTCCGCTACTACTAGCTA 367
            |||||||| | ||||||||||||| ||||||||||||| || ||||||||| |||
Db   69 CACACGGCGTCTGCCCAGAGAGAGCTGTGACCACCATTGCCACTACTATCTA 119

Db
LOCUS   BY783538   310 bp   mRNA   linear   EST 10-DEC-2003
DEFINITION BY783538 RIKEN full-length enriched, 17.5 days embryo whole body
            Mus musculus cDNA clone L930176D05 5', mRNA sequence.
ACCESSION   BY783538
VERSION   BY783538.1 GI:39710177
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 310)
  AUTHORS   Carninci,P., Waki,K., Shiraki,T., Konno,H., Shibata,K., Itoh,M.,
            Aizawa,K., Arakawa,T., Ishii,Y., Sasaki,D., Bono,H., Kondo,S.,
            Sugahara,Y., Saito,R., Osato,N., Fukuda,S., Sato,K., Watahiki,A.,
            Hirozane-Kishikawa,T., Nakamura,M., Shibata,Y., Yasunishi,A.,
            Kikuchi,N., Yoshiki,A., Kusakabe,M., Gustincich,S., Beisel,K.,
            Pavan,W., Aidinis,V., Nakagawara,A., Held,W.A., Iwata,H., Kono,T.,
            Nakauchi,H., Lyons,P., Wells,C., Hume,D.A., Fagiolini,M.,
            Hensch,T.K., Brinkmeier,M., Camper,S., Hirota,J., Mombaerts,P.,
            Muramatsu,M., Okazaki,Y., Kawai,J. and Hayashizaki,Y.
  TITLE   Targeting a complex transcriptome: the construction of the mouse
            full-length cDNA encyclopedia
  JOURNAL   Genome Res. 13 (6B), 1273-1289 (2003)
  MEDLINE   22703353
   PUBMED   12819125
COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp;
            URL:http://genome.gsc.riken.go.jp/
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in Riken
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site (http://genome.gsc.riken.jp/) for
            further details.
FEATURES   Location/Qualifiers
     source   1. .310

ORIGIN
                  /organism="Mus musculus"
                  /mol_type="mRNA"
                  /strain="C57BL/6J"
                  /db_xref="taxon:10090"
                  /clone="L930176D05"
                  /tissue_type="whole body"
                  /dev_stage=17.5 days embryo"
                  /clone_lib="RIKEN full-length enriched, 17.5 days embryo
                  whole body"

RESULT 3
BY794229

Query Match   Best Local Similarity   92.9%;   Score 87.8;   DB 13;   Length 310;
                                               Pred. No. 3.9e-15;
Matches 92;   Conservative   0;   Mismatches   7;   Indels   0;   Gaps   0;

Qy 269 ATGGTCCACTGCCCGCCGGAGACAACAAAACCCAGCGAGCCATTGAACACTGCACACGGCCATC 328
            ||||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||| |
Db   2 ATGGTCCACTGCCCGCGGAGACAACACACCCAGCGAGCCATTGAACACTGCACACGGCCGTC 61

Qy 329 TGCCCAGAGAGAGCTGTGACCACCACTTCCGCTACTACTA 367
            ||||||||||||||||||||||||| || |||||||||
Db   62 TGCCCAGAGAGAGCTGTGACCACCATTGCCACTACTATCTA 100

Db
LOCUS   BY794229   331 bp   mRNA   linear   EST 10-DEC-2003
DEFINITION BY794229 RIKEN full-length enriched, 17.5 days embryo whole body
            Mus musculus cDNA clone L930292H24 5', mRNA sequence.
ACCESSION   BY794229
VERSION   BY794229.1 GI:39720868
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 331)
  AUTHORS   Carninci,P., Waki,K., Shiraki,T., Konno,H., Shibata,K., Itoh,M.,
            Aizawa,K., Arakawa,T., Ishii,Y., Sasaki,D., Bono,H., Kondo,S.,
            Sugahara,Y., Saito,R., Osato,N., Fukuda,S., Sato,K., Watahiki,A.,
            Hirozane-Kishikawa,T., Nakamura,M., Shibata,Y., Yasunishi,A.,
            Kikuchi,N., Yoshiki,A., Kusakabe,M., Gustincich,S., Beisel,K.,
            Pavan,W., Aidinis,V., Nakagawara,A., Held,W.A., Iwata,H., Kono,T.,
            Nakauchi,H., Lyons,P., Wells,C., Hume,D.A., Fagiolini,M.,
            Hensch,T.K., Brinkmeier,M., Camper,S., Hirota,J., Mombaerts,P.,
            Muramatsu,M., Okazaki,Y., Kawai,J. and Hayashizaki,Y.
  TITLE   Targeting a complex transcriptome: the construction of the mouse
            full-length cDNA encyclopedia
  JOURNAL   Genome Res. 13 (6B), 1273-1289 (2003)
  MEDLINE   22703353
   PUBMED   12819125
COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp;
            URL:http://genome.gsc.riken.go.jp/
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in Riken
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site (http://genome.gsc.riken.jp/) for
            further details.
FEATURES   Location/Qualifiers
     source   1. .331
                  /organism="Mus musculus"
                  /mol_type="mRNA"
                  /strain="C57BL/6J"
                  /db_xref="taxon:10090"

RESULT 4

ORIGIN
```
        /clone="L930292H24"
        /tissue_type="whole body"
        /dev_stage="17.5 days embryo"
        /clone_lib="RIKEN full-length enriched, 17.5 days embryo
        whole body"
```

Query Match
```
Best Local Similarity   23.6%;    Score 87.8; DB 13;  Length 331;
Matches 92;  Conservative  0; Mismatches  7;  Indels  0;  Gaps  0;
```

```
Qy    269 ATGGTTCCACTGCCCGCCCGGAGACACAAAACCCAGCGAGCGAGCCATTGAACACTGCACACGGCGCCATC 328
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db      2 ATGGTTCCACTGCCCGCCCGGAGACACACCCAGCGAGCGAGCCATTGAACACTGCACACGGCGCCGTC  61

Qy    329 TGCCCAGAGAGCGTGTGACCACCACTTCCGCTACTRAGCTA 367
          ||||||||||||||||||||||||||||||||||||||||||
Db     62 TGCCCAGAGAGCGTGTGACCACCATTGCCACTACTRATCTA 100
```

LOCUS   BY771317   354 bp   mRNA   linear   EST 10-DEC-2003
DEFINITION BY771317 RIKEN full-length enriched, 17.5 days embryo whole body
            Mus musculus cDNA clone L930052G15 5', mRNA sequence.
ACCESSION   BY771317
VERSION   BY771317.1 GI:39697955
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 354)
  AUTHORS   Carninci,P., Waki,K., Shiraki,T., Konno,H., Shibata,K., Itoh,M.,
            Aizawa,K., Arakawa,T., Ishii,Y., Sasaki,D., Bono,H., Kondo,S.,
            Sugahara,Y., Saito,R., Osato,N., Fukuda,S., Sato,K., Watahiki,A.,
            Hirozane-Kishikawa,T., Nakamura,M., Shibata,Y., Yasunishi,A.,
            Kikuchi,N., Yoshiki,A., Kusakabe,M., Gustincich,S., Beisel,K.,
            Pavan,W., Aidinis,V., Nakagawara,A., Held,W.A., Iwata,H., Kono,T.,
            Nakauchi,H., Lyons,P., Wells,C., Hume,D.A., Fagiolini,M.,
            Hensch,T.K., Brinkmeier,M., Camper,S., Hirota,J., Mombaerts,P.,
            Muramatsu,M., Okazaki,Y., Kawai,J. and Hayashizaki,Y.
  TITLE   Targeting a complex transcriptome: the construction of the mouse
            full-length cDNA encyclopedia
  JOURNAL   Genome Res. 13 (6B), 1273-1289 (2003)
   MEDLINE
    PUBMED 12819125
  COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp;
            URL:http://genome.gsc.riken.go.jp/;
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site (http://genome.gsc.riken.jp/) for
            further details.
FEATURES   Location/Qualifiers
     source   1..354
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="L930052G15"
                     /tissue_type="whole body"
                     /dev_stage="17.5 days embryo"
                     /clone_lib="RIKEN full-length enriched, 17.5 days embryo
```

RESULT 5

ORIGIN                whole body"

Query Match
```
Best Local Similarity   23.6%;   Score 87.8; DB 13; Length 354;
Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
Qy    269 ATGGTTCCACTGCCCGCCCGGAGACACAAAACCCAGCGAGCGAGCCATTGAACACTGCACACGGCGCCATC 328
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db      2 ATGGTTCCACTGCCCGCCCGGAGACACACCCAGCGAGCGAGCCATTGAACACTGCACACGGCGCCGTC  61

Qy    329 TGCCCAGAGAGCGTGTGACCACCACTTCCGCTACTRAGCTA 367
          ||||||||||||||||||||||||||||||||||||||||||
Db     62 TGCCCAGAGAGCGTGTGACCACCATTGCCACTACTRATCTA 100
```

LOCUS   AI785818   401 bp   mRNA   linear   EST 02-JUL-1999
DEFINITION ui78h05.y1 Sugano mouse liver mlia Mus musculus cDNA clone
            IMAGE:1888569 5' similar to gb:X81579 M.musculus mRNA for
            insulin-like growth factor binding (MOUSE);, mRNA sequence.
ACCESSION   AI785818
VERSION   AI785818.1 GI:5333534
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 401)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE   The WashU-NCI Mouse EST Project
  JOURNAL   Unpublished (1999)
  COMMENT   Other ESTs: ui78h05.x1
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: custom primer used
            High quality sequence stop: 126.
            MGI:972893
FEATURES   Location/Qualifiers
     source   1..401
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1888569"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse liver mlia"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCGGCCGCTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCCTACTCG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). Xho1 should
                     be used to isolate the cDNA insert. Size selection was
                     performed to exclude fragments <1.5kb. Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science). Custom primers for
                     sequencing: 5' end primer CTTCTCGTCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
```

RESULT 6

ORIGIN
```
Query Match    23.3%;   Score 86.8;  DB 9;  Length 401;
Best Local Similarity  92.9%;  Pred. No. 8.7e-15;
Matches  91;  Conservative  0;  Mismatches  7;  Indels  0;  Gaps  0;

Qy  270 TGGTCCACTGCCCGCGCCGAGACAAACCCAGCGAGCAATTGAACACTGCACACGGCCATCT 329
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    1 TGGTCCACTGCCCGCGCCGAGACAAACCCAGCGAGCAATTGAACACTGCACACGGCCGTCT 60

Qy  330 GCCCAGAGAGCTGTGACCACCACTTCCGCTACTAGCTA 367
        ||||||||||||||||||||||||||| ||||||||||
Db   61 GCCCAGAGAGCTGTGACCACCACCATTGCCACTACTATCTA 98
```

```
LOCUS       AI196314      480 bp    mRNA    linear   EST 14-OCT-1998
DEFINITION  ui71a07.y1 Sugano mouse liver mlia Mus musculus cDNA clone
            IMAGE:1887828 5', similar to gb:M59316 rnal INSULIN-LIKE GROWTH
            FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
            mRNA for insulin-like growth factor binding (MOUSE);, mRNA
            sequence.
ACCESSION   AI196314
VERSION     AI196314.1  GI:3748920
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 480)
AUTHORS     Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE       The WashU-HHMI Mouse EST Project
JOURNAL     Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:972152
            Seq primer: custom primer used
            High quality sequence stop: 375.
FEATURES             Location/Qualifiers
     source          1..480
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1887828"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse liver mlia"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCCTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert. Size selection was
                     performed to exclude fragments <1.5kb. Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science). Custom primers for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTCG and 3' end
```

RESULT 7

ORIGIN   primer CGACCTGCAGCTCGAGCACA. "
```
Query Match    23.3%;   Score 86.8;  DB 9;  Length 480;
Best Local Similarity  92.9%;  Pred. No. 9.5e-15;
Matches  91;  Conservative  0;  Mismatches  7;  Indels  0;  Gaps  0;

Qy  270 TGGTCCACTGCCCGCGCCGAGACAAACCCAGCGAGCAATTGAACACTGCACACGGCCATCT 329
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    1 TGGTCCACTGCCCGCGCCGAGACAAACCCAGCGAGCAATTGAACACTGCACACGGCCGTCT 60

Qy  330 GCCCAGAGAGCTGTGACCACCACTTCCGCTACTAGCTA 367
        ||||||||||||||||||||||||||| ||||||||||
Db   61 GCCCAGAGAGCTGTGACCACCACCATTGCCACTACTATCTA 98
```

```
LOCUS       AI530146      706 bp    mRNA    linear   EST 18-MAR-1999
DEFINITION  ui89f09.y1 Sugano mouse liver mlia Mus musculus cDNA clone
            IMAGE:1889609 5', similar to gb:M59316 rnal INSULIN-LIKE GROWTH
            FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
            mRNA for insulin-like growth factor binding (MOUSE);, mRNA
            sequence.
ACCESSION   AI530146
VERSION     AI530146.1  GI:4442281
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 706)
AUTHORS     Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
TITLE       The WashU-NCI Mouse EST Project 1999
JOURNAL     Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:973933
            Seq primer: custom primer used
            High quality sequence stop: 479.
FEATURES             Location/Qualifiers
     source          1..706
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1889609"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse liver mlia"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCCTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert. Size selection was
                     performed to exclude fragments <1.5kb. Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science). Custom primers for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTCG and 3' end
```

ORIGIN

primer CGACCTGCAGCTCGAGCACA."

RESULT 8
Query Match   23.3%;   Score 86.8; DB 9;   Length 706;
Best Local Similarity 92.9%;   Pred. No. 1.2e-14;
Matches 91; Conservative   0; Mismatches 7; Indels   0; Gaps   0;

Qy   270 TGGTCCACTGCCCCCGGGAGACACAAAACCCAGGAGGCATTGAAACACTGCACGGCCCATCT 329
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TGGTCCACTGCCCCCGGGAGACACACAACCCAGGCGAGGCATTGAACACTGCACACGGGCCGTCT 60

Qy   330 GCCCAGAGAGGCTGTGACCACCACTTCCGGTACTTAGCTA 367
          |||||||||||||||||||| ||| ||||||| |||||||
Db   61 GCCCAGAGAGGCTGTGACCACCACCATTGCCACTACTACTA 98

ACCESSION   AI098594
VERSION   AI098594.1 GI:3448119
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 765)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE   The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT   Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:930344
            Seq primer: custom primer used
            High quality sequence stop: 395.
FEATURES   Location/Qualifiers
     source   1..765
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1481988"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse liver mlia"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCCTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert. Size selection was
                     performed to exclude fragments <1.5kb. Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science). Custom primers for
                     sequencing: 5' end primer CTTCTCGCTCTAAAAGTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."

RESULT 9
Query Match   23.3%;   Score 86.8; DB 9;   Length 765;
Best Local Similarity 92.9%;   Pred. No. 1.2e-14;
Matches 91; Conservative   0; Mismatches 7; Indels   0; Gaps   0;

Qy   270 TGGTCCACTGCCCCCGGGAGACACAAAACCCAGGAGGCATTGAAACACTGCACGGCCCATCT 329
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TGGTCCACTGCCCCCGGGAGACACACAACCCAGGCGAGGCATTGAACACTGCACACGGGCCGTCT 60

Qy   330 GCCCAGAGAGGCTGTGACCACCACTTCCGGTACTTAGCTA 367
          |||||||||||||||||||| ||| ||||||| |||||||
Db   61 GCCCAGAGAGGCTGTGACCACCACCATTGCCACTACTACTA 98

ACCESSION   AI530313
VERSION   AI530313.1 GI:4444448
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 785)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE   The WashU-NCI Mouse EST Project
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:974109
            Seq primer: custom primer used
            High quality sequence stop: 459.
FEATURES   Location/Qualifiers
     source   1..785
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1889785"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse liver mlia"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCCTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert. Size selection was
                     performed to exclude fragments <1.5kb. Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science). Custom primers for
                     sequencing: 5' end primer CTTCTCGCTCTAAAAGTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."

DEFINITION (RESULT 8) AI098594 ue31e07.y1 Sugano mouse liver mlia Mus musculus cDNA clone IMAGE:1481988 5' similar to gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.   765 bp mRNA   linear EST 20-AUG-1998

DEFINITION (RESULT 9) AI530313 ui91f01.y1 Sugano mouse liver mlia Mus musculus cDNA clone IMAGE:1889785 5' similar to gb:M59316_rna1 INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.   785 bp mRNA   linear EST 18-MAR-1999

RESULT 10

```
Query Match    23.3%;  Score 86.8;  DB 9;  Length 785;
Best Local Similarity   92.9%;  Pred. No. 1.2e-14;
Matches   91;  Conservative   0;  Mismatches   7;  Indels   0;  Gaps   0;

Qy   270 TGGTCCACTGCCCGCCGAGACAACAAACCCAGCGAGCATTGAACACTGCACGGCCATCT 329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     1 TGGTCCACTGCCCGCCGAGACAACACCCCAGCGAGCATTGAACACTGCACACGGCCGTCT  60

Qy   330 GCCCAGAGAGCTGTGACCACCACTTCCGCTACTAGCTA 367
             ||||||||||||||||||||||||||||||| |||||||
Db    61 GCCCAGAGAGCTGTGACCACCATTGCCACTACTATCTA  98
```

ORIGIN

LOCUS   AI529939   799 bp   mRNA   linear   EST 18-MAR-1999
DEFINITION ui87c09.y1 Sugano mouse liver mlia Mus musculus cDNA clone
            IMAGE:1889392 5' similar to gb:M59316 rna1 INSULIN-LIKE GROWTH
            FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
            mRNA for insulin-like growth factor binding (MOUSE);, mRNA
            sequence.
ACCESSION   AI529939
VERSION   AI529939.1 GI:4444074
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 799)
AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
TITLE   The WashU-NCI Mouse EST Project 1999
JOURNAL   Unpublished (1999)
COMMENT   Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:973716
            Seq primer: custom primer used
            High quality sequence stop: 506.
FEATURES   Location/Qualifiers
     source   1..799
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1889392"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse liver mlia"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert. Size selection was
                     performed to exclude fragments <1.5kb. Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science). Custom primers for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
ORIGIN
```

RESULT 11

```
Query Match    23.3%;  Score 86.8;  DB 9;  Length 799;
Best Local Similarity   92.9%;  Pred. No. 1.2e-14;
Matches   91;  Conservative   0;  Mismatches   7;  Indels   0;  Gaps   0;

Qy   270 TGGTCCACTGCCCGCCGAGACAACAAACCCAGCGAGCATTGAACACTGCACGGCCATCT 329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     1 TGGTCCACTGCCCGCCGAGACAACACCCCAGCGAGCATTGAACACTGCACACGGCCGTCT  60

Qy   330 GCCCAGAGAGCTGTGACCACCACTTCCGCTACTAGCTA 367
             ||||||||||||||||||||||||||||||| |||||||
Db    61 GCCCAGAGAGCTGTGACCACCATTGCCACTACTATCTA  98
```

ORIGIN

LOCUS   AI790802   848 bp   mRNA   linear   EST 02-JUL-1999
DEFINITION uk28b10.y1 Sugano mouse kidney mkia Mus musculus cDNA clone
            IMAGE:1970299 5' similar to gb:X81579 M.musculus mRNA for
            insulin-like growth factor binding (MOUSE);, mRNA sequence.
ACCESSION   AI790802
VERSION   AI790802.1 GI:5338518
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 848)
AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
TITLE   The WashU-NCI Mouse EST Project 1999
JOURNAL   Unpublished (1999)
COMMENT   Other ESTs: uk28b10.x1
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:987039
            Seq primer: custom primer used
            High quality sequence stop: 514.
FEATURES   Location/Qualifiers
     source   1..848
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1970299"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse kidney mkia"
                     /note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert. Size selection was
                     performed to exclude fragments <1.5kb. Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science). Custom primers for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
ORIGIN
```

RESULT 12

```
Query Match     23.3%;  Score 86.8; DB 9;  Length 848;
Best Local Similarity  92.9%;  Pred. No. 1.3e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy  270 TGGTCCACTGCCCGCCGGAGACAAAACCCAGCGAGCGCATTGAACACTGCACACGGCGCCATCT 329
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    1 TGGTCCACTGCCCGCCGGAGACACACCCAGCGAGCGCATTGAACACTGCACACGGCGCCGTCT 60

Qy  330 GCCCAGAGAGAGCTGTGACCACCACTTCCGCTACTAGCTA 367
        |||||||||||||||||||||||||||||||| ||||||||
Db   61 GCCCAGAGAGAGCTGTGACCACCACTTGCCACTACTACTA 98
```

LOCUS   AI528304   852 bp   mRNA   linear   EST 18-MAR-1999
DEFINITION ui95g10.y1 Sugano mouse liver m1a Mus musculus cDNA clone
            IMAGE:1890210 5' similar to gb:M59316 rnal INSULIN-LIKE GROWTH
            FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
            mRNA for insulin-like growth factor binding (MOUSE);;, mRNA
            sequence.
ACCESSION   AI528304
VERSION   AI528304.1 GI:4442439
KEYWORDS   EST.
SOURCE   Mus musculus
  ORGANISM Mus musculus (house mouse)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 852)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE   The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Marra M/WashU-NCI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:974534
            Seq primer: custom primer used
            High quality sequence stop: 478.
FEATURES   Location/Qualifiers
     source   1..852
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1890210"
                     /sex="female"
                     /dev_stage="adult"
                     /clone_lib="Sugano mouse liver m1a"
                     /lab_host="DH10B"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert. Size selection was
                     performed to exclude fragments <1.5kb. Library
                     constructed by Dr. Sunio Sugano (University of Tokyo
                     Institute of Medical Science). Custom primers for
                     sequencing: 5' end primer CTTCTGCCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
ORIGIN
```

RESULT 13

```
Query Match     23.3%;  Score 86.8; DB 9;  Length 852;
Best Local Similarity  92.9%;  Pred. No. 1.3e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy  270 TGGTCCACTGCCCGCCGGAGACAAAACCCAGCGAGCGCATTGAACACTGCACACGGCGCCATCT 329
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    1 TGGTCCACTGCCCGCCGGAGACACACCCAGCGAGCGCATTGAACACTGCACACGGCGCCGTCT 60

Qy  330 GCCCAGAGAGAGCTGTGACCACCACTTCCGCTACTAGCTA 367
        |||||||||||||||||||||||||||||||| ||||||||
Db   61 GCCCAGAGAGAGCTGTGACCACCACTTGCCACTACTACTA 98
```

LOCUS   AI196154   605 bp   mRNA   linear   EST 14-OCT-1998
DEFINITION ui69d08.y1 Sugano mouse liver m1a Mus musculus cDNA clone
            IMAGE:1887663 5' similar to gb:M59316 rnal INSULIN-LIKE GROWTH
            FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
            mRNA for insulin-like growth factor binding (MOUSE);;, mRNA
            sequence.
ACCESSION   AI196154
VERSION   AI196154.1 GI:3748760
KEYWORDS   EST.
SOURCE   Mus musculus
  ORGANISM Mus musculus (house mouse)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 605)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE   The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT   Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:971987
            Seq primer: custom primer used
            High quality sequence stop: 420.
FEATURES   Location/Qualifiers
     source   1..605
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1887663"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse liver m1a"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert. Size selection was
                     performed to exclude fragments <1.5kb. Library
                     constructed by Dr. Sunio Sugano (University of Tokyo
                     Institute of Medical Science). Custom primers for
                     sequencing: 5' end primer CTTCTGCCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
```

ORIGIN

Query Match   22.9%;   Score 85.2;   DB 9;   Length 605;
Best Local Similarity   91.8%;   Pred. No. 3.2e-14;
Matches 90; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy   270 TGGTCCACTGCCCGCGCCGAGACAACAACCCAGGCAGCATTGAACACTGCACGGCCATCT 329
         |||||||||||||||||||||||||||| ||||||||||||||||||||||||||| |||
Db   1 TGGTCCACTGCCCGCGCCGAGACAACACACCCAGGCAGCATTGAACACTGCACGGTCGTCT 60

Qy   330 GCCCAGAGAGCTGTGACCACCACTTCCGCTACTAGCTA 367
         |||||||||||||||||||||||||||||||||| |||
Db   61 GCCCAGAGAGCTGTGACCACCACTTCCGCTACTATCTA 98

RESULT 14
LOCUS   AI785039   380 bp   mRNA   linear   EST 02-JUL-1999
DEFINITION ui73a06.y1 Sugano mouse liver mlia Mus musculus cDNA clone
            IMAGE:1888018 5' similar to gb:X81579 M.musculus mRNA for
            insulin-like growth factor binding (MOUSE);, mRNA sequence.
ACCESSION   AI785039
VERSION   AI785039.1 GI:5332755
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 380)
AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
TITLE   The WashU-NCI Mouse EST Project 1999
JOURNAL   Unpublished (1999)
COMMENT   Other_ESTs: ui73a06.x1
            Contact: Marra M/WashU-NCI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1810
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: custom primer used
            High quality sequence stop: 353.
FEATURES   Location/Qualifiers
     source   1..380
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1888018"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse liver mlia"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCGTTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert. Size selection was
                     performed to exclude fragments <1.5kb. Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science). Custom primers for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTCGG and 3' end
                     primer CGACCTGCAGCTCGAGAGCACA."
ORIGIN

Query Match   22.1%;   Score 82.2;   DB 9;   Length 380;
Best Local Similarity   91.6%;   Pred. No. 2e-13;
Matches 87; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy   273 TCCACTGCCCGCGCCGAGACAACAAACCCAGGCAGCGAGCATTGAACACTGCACGGTCGTCTGCC 332
         |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   1 TCCACTGCCCGCGCCGAGACAACACACCCAGGCAGCGAGCATTGAACACTGCACGGTCGTCTGCC 60

Qy   333 CAGAGAGCTGTGACCACCACTTCCGCTACTAGCTA 367
         ||||||||||||||||||||||||||| |||
Db   61 CAGAGAGCTGTGACCACCACTTCCGCTACTATCTA 95

RESULT 15
LOCUS   W30013   332 bp   mRNA   linear   EST 11-SEP-1996
DEFINITION mc24c07.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
            IMAGE:349452 5' similar to gb:X81579 M.musculus mRNA for
            insulin-like growth factor binding (MOUSE);, mRNA sequence.
ACCESSION   W30013
VERSION   W30013.1 GI:1310178
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 332)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   Contact: Marra M/Mouse EST Project
            Washington-HHMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1810
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: mob.REGA+ET
            High quality sequence stop: 250.
FEATURES   Location/Qualifiers
     source   1..332
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:349452"
                     /dev_stage="19.5 dpc total fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares mouse p3NMF19.5"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGAGCGGCCGCGCATTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA was kindly provided by
                     Dr. Minoru Ko (Wayne State University)."
ORIGIN

Query Match   19.7%;   Score 73.2;   DB 14;   Length 332;
Best Local Similarity   90.7%;   Pred. No. 8.5e-11;
Matches 78; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy    282 CGCCGAGACACACAAACCCAGCGAGCGCATTGAAACACTGCACACGGCGCATCTGCCCAGAGAGCT 341
          |  ||||||||||| ||||||||||||||||||||||||| ||||||| |||||||||||||||||
Db     17 CGAGGAGACACACACCCAGCGAGCCATTGAACACTGCACACGGGCCGTCTGCCCAGAGAGCT 76

Qy    342 GTGACCACCACACTTCCGCTRACTAGCTA 367
          |||||||||||| |  |  |||||| |||
Db     77 GTGACCACCATTGCCACTACTACTATCTA 102
```

Search completed: June 7, 2004, 14:45:58
Job time : 2495 secs